THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XXXVII November, 1946 Number 8

A COMPARISON OF THE PRINCIPAL-AXIS AND CENTROID FACTORS

KARL J. HOLZINGER
The University of Chicago

1. INTRODUCTION

The object of the present paper is chiefly to present a comparison of the principal-axis and centroid methods when variances and covariances are analyzed instead of the usual unit variances and correlations. One justification for such a paper is in the confusion often arising in the minds of psychologists and even some factor analysts in interpreting the covariance form of solution as presented by Kelley[1] or Davis.[2] No mathematical criticism of such forms of presentation is either warranted or intended, but the writer feels that if these forms are expressed in a different way and linked with the usual type of solution some semantic advantage may be obtained. He also believes that conceptual learning is greatly implemented by looking at a problem from a number of angles some combination of which may finally lead the student to say: "Oh, now I see more clearly what this is all about." Since writing this paper helped the writer himself somewhat in this manner, it is hoped it will be of some value to other students of factor analysis.

An arbitrary two-variable problem (twelve cases) has been selected so that all relationships may be indicated in two dimensions. Two artificial variables are ample for exhibiting all the essential differences in method, inasmuch as they are of necessity mathematical, and most debatable points of difference are

[1] Truman L. Kelley, *Essential Traits of Mental Life.* Cambridge, Mass.: Harvard University Press, 1935.

[2] Frederick B. Davis, "Fundamental Factors of Comprehension in Reading," *Psychometrika*, IX (September, 1944), 185–97.

revealed as clearly by two artificial tests as by twenty observed variables.

One point to be disposed of at the outset is that of 'communalities.' By factoring total variances and covariances we are making no separate allowance for usual factors called specificity or unreliability. We are simply factoring the whole of the tests just as we do when we insert unities in the diagonal of a correlation matrix. Whether or not it is wiser to factor the whole test or the communality is of no concern in the present paper because it has no bearing upon the comparisons of methods here made. Although the writer has factored communalities almost exclusively for twenty-five years, he would certainly be receptive to the use of total variances for some types of problems, and not necessarily rule out the latter type of method by the usual meaningless psychological jargon of 'psychologically meaningless solutions.'

Since the differences in the two methods are really mathematical, mathematics cannot be avoided, but the writer would have liked to make the presentation entirely in nonmathematical language if it were possible to clarify the discussion in this way. As a compromise, as simple mathematics as possible has been used to present the ideas as the writer views them, and a considerable amount of interpretation of what the mathematics mean has been attempted. Even if the reader has not had the slight amount of trigonometry and calculus here used, he may yet follow the main argument (we hope). Two new formulas (but of course no new method) are derived in vector representation for an alternate interpretation of methods. The example used is for illustration only and for practical purposes is of little value because the factoring of two tests is usually trivial. It will also be obvious that we are not at all interested here in the general routine of factor analysis for several variables, but only in the meaning of the methods as revealed by factoring two tests.

We shall also refrain from comment in this paper on Kelley's interpretation of the centroid solution made over ten years ago except to say that we disagree entirely with his statement on page 62 that the centroid has "the weakness of an arithmetic average of semi-disparate things," while the principal axis is "interpretable." We believe both methods are interpretable, that both are not only averages of semi-disparate things but of

very disparate things, and that it does not matter because all statistics are of that nature even though we try to imagine otherwise.

2. FACTOR PATTERNS AND TRANSFORMATIONS

In the following discussion tests or factors are expressed either in deviate or in standard form. The tests will be denoted by x and y, with standard deviations σ_x and σ_y. If $z_x = x/\sigma_x$ and $z_y = y/\sigma_y$ denote the standard form of these variables, the following varieties and covariances may be given: $v_x = \sigma_x^2$, $v_y = \sigma_y^2$, $v_{z_x} = 1$, $v_{z_y} = 1$, $p_{xy} = \Sigma xy/N = r_{xy}\sigma_x\sigma_y$, $p_{z_xz_y} = \Sigma_{z_xz_y}/N = r_{xy}$. The principal-axis factors are denoted by p_A and p_B, while the centroid factors are designated as c_a and c_b. In the general discussion x_1 and x_2 will denote any pair of factors. It will be assumed in the present development that r_{xy} is positive and that the correlation between any of the above pairs of factors is zero. A summary of the above notation is given for the convenience of the reader.

SUMMARY OF NOTATION

Variables	Deviate Form	Standard Form	Standard Deviation	Variance
Tests	x, y	z_x, z_y	σ_x, σ_y	v_x, v_y
General type factors	x_1, x_2	z_1, z_2	σ_1, σ_2	v_1, v_2
Principal factors	p_A, p_B	z_A, z_B	σ_A, σ_B	v_A, v_B
Centroid factors	c_a, c_b	z_a, z_b	σ_a, σ_b	v_a, v_b

Our present problem is to express the tests or variables x and y in terms of the uncorrelated factors x_1 and x_2. Inasmuch as linear relationships are the simplest mathematically and lend themselves to simple statistical treatment, they are the forms employed in factor analysis. These linear expressions, which we shall presently write, are therefore adopted for mathematical and statistical convenience. Their arbitrary nature is stressed at this point because it is a characteristic of all descriptive science which 'discovers' nothing, but attempts to explain things more simply and exactly than do informal methods. With these reservations in mind we may write the following two equations:

$$\left.\begin{aligned} x &= a_1x_1 + a_2x_2 \\ y &= b_1x_1 + b_2x_2 \end{aligned}\right\} \begin{matrix}\text{Factor pattern} \\ \text{with all variables} \\ \text{in deviate form}\end{matrix} \qquad (1)$$

where a_1, a_2, b_1, and b_2 are numerical coefficients. These equations are often called a 'factor pattern' or 'solution'. In mathematics they might be called a linear transformation expressing the variables x and y in terms of the factors x_1 and x_2. Equations (1) will presently be expressed in trigonometric form for clarity of geometric interpretation.

In the factor analysis we are obviously concerned with the determination of the four coefficients in equations (1). Once these are known, the solution or factor pattern is complete, giving an exact picture of two tests in terms of uncorrelated factors. We are also concerned in factor analysis with the

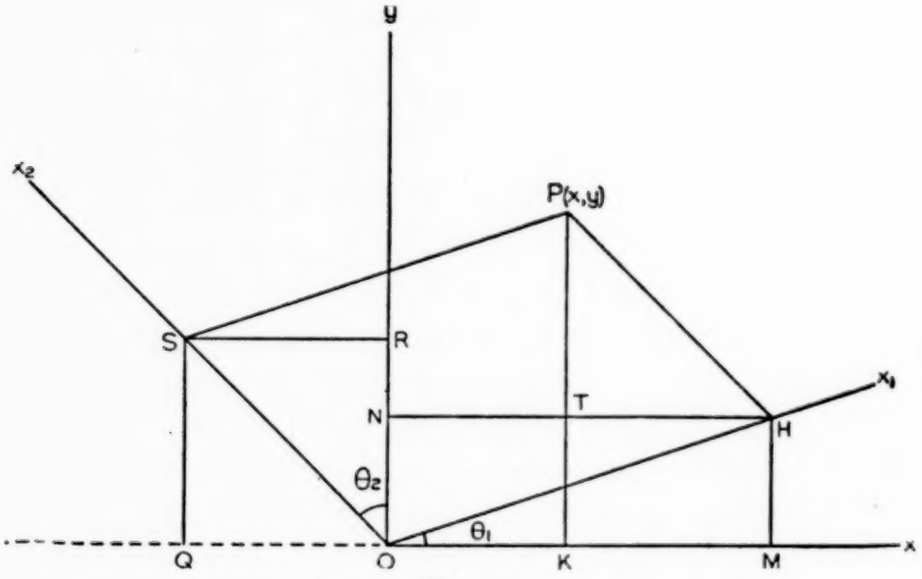

FIG. 1.

determination of the factors themselves. These may be obtained by solving equations (1) for x_1 and x_2, giving expressions of the form

$$\left.\begin{aligned} x_1 &= A_1x + B_1y \\ x_2 &= A_2x + B_2y \end{aligned}\right\} \begin{array}{l}\text{Transformation expressing } x_1 \text{ and } x_2 \text{ in terms} \\ \text{of } x \text{ and } y\end{array} \qquad (2)$$

We shall next present a simple development of transformation (1) in trigonometric form. Let the test axes be Ox and Oy and factor axes be Ox_1 and Ox_2 rotated through angles θ_1 and θ_2, as shown in Fig. 1. Different angles are assumed in order to indicate the most general case. The coördinates of any point $P(x,y)$ will now be expressed in terms of x_1 and x_2. In Fig. 1

HP is drawn parallel to OS, and SP is parallel to OH. All other lines which appear to be either perpendicular or parallel may be taken as such without further specification. It therefore follows that $OH = x_1$, $OS = x_2$, and $x = OK = OM - KM = OM - TH = OM - OQ$.[1] But $OM = OH \cos \theta_1 = x_1 \cos \theta_1$, and $OQ = OS \sin \theta_2 = x_2 \sin \theta_2$. Therefore, $x = x_1 \cos \theta_1 - x_2 \sin \theta_2$. It is also apparent from Fig. 1 that $y = KP = ON + TP = ON + OR = OH \sin \theta_1 + OS \cos \theta_2 = x_1 \sin \theta_1 + x_2 \cos \theta_2$. The required general transformation is therefore:

$$\left.\begin{aligned} x &= x_1 \cos \theta_1 - x_2 \sin \theta_2 \\ y &= x_1 \sin \theta_1 + x_2 \cos \theta_2 \end{aligned}\right\} \begin{matrix}\text{General transformation} \\ \text{expressing } x \text{ and } y \\ \text{in terms of } x_1 \text{ and } x_2\end{matrix} \qquad (3)$$

It will be shown later that the centroid is an example of the general form (3). The principal axis solution, however, is one for which $\theta_1 = \theta_2 = \theta$. In the latter special case of equations (3) we have what is known as an orthogonal transformation of the form

$$\left.\begin{aligned} x &= x_1 \cos \theta - x_2 \sin \theta \\ &= p_A \cos \theta - p_B \sin \theta \\ y &= x_1 \sin \theta + x_2 \cos \theta \\ &= p_A \sin \theta + p_B \cos \theta \end{aligned}\right\} \begin{matrix}\text{Orthogonal transforma-} \\ \text{tion expressing } x \text{ and } y \\ \text{in terms of } x_1 \text{ and } x_2\end{matrix} \qquad (4)$$

Solving for x_1 and x_2 gives another orthogonal transformation:

$$\left.\begin{aligned} x_1 = p_A &= \quad x \cos \theta + y \sin \theta \\ x_2 = p_B &= -x \sin \theta + y \cos \theta \end{aligned}\right\} \begin{matrix}\text{Orthogonal transforma-} \\ \text{tion expressing } x_1 \text{ and } x_2 \\ \text{in terms of } x \text{ and } y\end{matrix} \qquad (5)$$

It will be noted that the sums of squares of the coefficients in (4) and (5) are equal to one both by rows and by columns, and that the sum of the pair-products is zero. These are properties of orthogonal transformations.

Equations (4) and (5) are the forms employed by Kelley in presenting his solutions. In addition to this part of the solution, he also gives the variances of p_A and p_B obtained from equations (5):[2]

[1] In the present case we are considering only the absolute magnitude of OQ.

[2] The complete algebra for the derivation of v_A will be given here in full but will not be repeated for similar expressions, which the student should

$$\begin{aligned} v_A &= v_x \cos^2 \theta + v_y \sin^2 \theta + 2p_{xy} \sin \theta \cos \theta \\ v_B &= v_x \sin^2 \theta + v_y \cos^2 \theta - 2p_{xy} \sin \theta \cos \theta. \end{aligned} \tag{6}$$

It is also apparent from equations (4) that

$$\begin{aligned} v_x &= v_A \cos^2 \theta + v_B \sin^2 \theta \\ v_y &= v_A \sin^2 \theta + v_B \cos^2 \theta \end{aligned} \tag{7}$$

$$v_x + v_y = v_A \qquad + v_B \tag{8}$$

An orthogonal pattern of form (4) bordered by appropriate variances would then appear in the form of Table 1. This is the

TABLE 1

	p_A	p_B	Variance
x	$\cos \theta$	$-\sin \theta$	v_x
y	$\sin \theta$	$\cos \theta$	v_y
Variance	v_A	v_B	Total

form employed by Davis and should be understood to be a combination of the factor pattern (4), the transformation (5), and the results of equations (6), (7), and (8).

It may be desirable at this point to introduce our illustrative example before discussing forms of solution alternate to Table 1. We shall therefore merely give the data and solutions, leaving the derivations of the latter to the next sections. The data [Table 3] are twelve artificial pairs of scores in deviate form so skillfully selected by Swineford that their correlation is exactly .72, $\sigma_x = 15$, $\sigma_y = 10$, $p_{xy} = 108$, and the distributions are approximately normal. All we need for the complete solution of equations (4) to (8), or Table 1, is the angle θ. This is found in Section 3 to be approximately 29°58′. Knowing this angle

---

learn to read immediately without the following type of analysis. Squaring both members of the first of equations (5) gives

$$p_A^2 = x^2 \cos^2 \theta + y^2 \sin^2 \theta + 2xy \cos \theta \sin \theta.$$

Summing from 1 to N and dividing by N yields

$$\frac{\Sigma p_A^2}{N} = \frac{\Sigma x^2}{N} \cos^2 \theta + \frac{\Sigma y^2}{N} \sin^2 \theta + \frac{2\Sigma xy}{N} \cos \theta \sin \theta,$$

or

$$v_A = v_x \cos^2 \theta + v_y \sin^2 \theta + 2p_{xy} \cos \theta \sin \theta.$$

we can immediately write the solution (4) from trigonometric tables as

$$\begin{aligned} x &= .8663p_A - .4995p_B \\ y &= .4995p_A + .8663p_B, \end{aligned} \qquad (4)'$$

and equations (5) give

$$\begin{aligned} p_A &= .8663x + .4995y \\ p_B &= -.4995x + .8663y. \end{aligned} \qquad (5)'$$

Equations (6) then yield the variances of the factors:

$$\begin{aligned} v_A &= .7505v_x + .2495v_y + .8654p_{xy} = 287.3 \\ v_B &= .2495v_x + .7505v_y - .8654p_{xy} = 37.7 \end{aligned} \qquad (6)'$$

$$v_A + v_B = \quad v_x + \quad v_y \quad = 325.0$$

or, in tabular form:

TABLE 1′

	p_A	p_B	Variance
x	.8663	−.4995	225
y	.4995	.8663	100
Variance	287.3	37.7	325.0

The values of the factors p_A and p_B may be found from equations (5)′ by substituting the necessary values of x and y from Table 3. The resulting factor scores for p_A and p_B are also shown in this table. The correlation of these twelve pairs of factor scores is, of course, equal to zero (within rounding error). Equations of the form (5)′ have sometimes been referred to as 'regression equations.'[1] This is obviously not an accurate description, since the latter imply estimation in the usual statistical meaning of such expressions. Equations (5)′ give exact determinations of the factors without the usual clumsy methods of estimation, and this property is one of the chief virtues of the principal-axis solution in the forms (4) and (5). To call equations (5)′ estimates is therefore a denial of one of the chief merits of the whole method (or any method involving as many factors as tests so that an algebraic solution of the latter is possible).

We are now ready to interpret the numerical principal-axis

[1] Davis, *loc. cit.*, p. 190.

solution given in Table 1′. The numbers in the body of the table, or from equations (4)′, are trigonometric functions in the orthogonal transformation, but we may interpret them as 'loadings' in the usual manner provided the variances of all variables are properly considered. Taking the second of equations (4)′ as an illustration, y appears to be more heavily 'saturated' with p_B than with p_A. This saturation is more clearly expressed by variance, using equations (7):

$$\begin{aligned} v_x &= .7505v_A + .2495v_B \\ &= .7505(287.3) + .2495(37.7) = 215.6 + 9.4 \\ v_y &= .2495v_A + .7505v_B \\ &= .2495(287.3) + .7505(37.7) = 71.7 + 28.3 \end{aligned} \qquad (7)'$$

We may therefore say that of the total variance, $v_y = 100$, 71.7 is 'due to' or 'attributable to' the factor p_A or, preferably, to the variance of p_A, while only 28.3 is due to p_B. These expressions will not be misleading provided we do not read into them anything in the nature of cause or directional influence. It would therefore probably be safer to say that the total variance of v_y is 'composed of' twenty-five per cent of the variance of v_A, or 71.7, plus seventy-five per cent of the variance of v_B, or 28.3. This last sentence is merely the English version of the second of equations (7)′.

Whatever language may be employed, the fact remains that the four numbers at the right of equations (7)′ give a basis for comparing the importance of factors. The ratio of 215.6 to 28.3, for example, is approximately 8, which means that p_A is eight times as important in the determination of v_x as is p_B in the determination of v_y. Looking back at Table 1′ we see that the two numbers in the body of the table corresponding to 215.6 and 28.3 are the same; namely, .8663. To interpret the importance of factors by looking at Table 1′, we must therefore actually perform in the mind the operations given in equations (7)′, or we might note that the entry .8663 is the same in this case and, therefore, judge the importance of p_A in determining v_x to that of p_B in determining v_y by the ratio of 287.3 ÷ 37.7, which is again approximately eight.

None of the foregoing remarks are a criticism of Kelley's form as an accurate representation, but merely as a form difficult for many students to interpret. The temptation to gaze amongst

the relatively meaningless direction cosines in the body of the table looking for importance of factors is almost irresistible. It was a desire to resist this urge that led to the present article.

In an attempt to expand the foregoing interpretation we shall next express the above solution in the form of standardized factors, which will bring σ_A and σ_B directly into the 'loadings' inspected for importance of factors. To accomplish this we merely need to substitute in equations (4) the values $p_A = \sigma_A z_A$ and $p_B = \sigma_B z_B$, where z_A and z_B are the factors p_A and p_B expressed in standard form. Making these substitutions we find

$$\begin{aligned} x &= (\sigma_A \cos\theta) z_A - (\sigma_B \sin\theta) z_B \\ y &= (\sigma_A \sin\theta) z_A + (\sigma_B \cos\theta) z_B \end{aligned} \tag{9}$$

or, in numerical form,

$$\begin{aligned} x &= 14.68 z_A - 3.07 z_B \\ y &= 8.47 z_A + 5.32 z_B \end{aligned} \tag{9'}$$

The variances from equations (9)′ are of course identical with those in equations (7)′, since the variances of z_A and z_B are equal to 1, but we shall recalculate them from equations (9)′ as a check:

$$\begin{aligned} v_x &= (14.68)^2 + (3.07)^2 = 215.5 + 9.4 = 224.9 \\ v_y &= (8.47)^2 + (5.32)^2 = 71.7 + 28.3 = 100.0 \end{aligned} \tag{10}$$

$$\begin{aligned} v_x + v_y &= 287.2 + 37.7 = 324.9 \\ &= v_A + v_B \end{aligned}$$

The recommended form of solution is (9)′ and (10), or a tabular form like Table 1′.

TABLE 2

	z_A	z_B	Variance
x	14.68	−3.07	224.9
y	8.47	5.32	100.0
Variance	287.2	37.7	324.9

The product-moment, p_{xy}, is equal to (14.68)(8.47) − (3.07)(5.32) = 108.

If we now look at equations (9)′ and (10), or Table 2, we see at once that p_A is much more important in determining v_x than

is p_B in determining v_y by observing that 14.68 is much larger than 5.32. The ratio of importance is shown by squaring these numbers as in equations (10). Another reason for employing equations of form (9) is that they may be linked with the usual interpretation of loadings for standardized tests. For example, in equations (9)′ we may say that of the variance, $v_x = 225$, 215.5 times the unit variance of z_A is attributable to p_A. The chief difference between this and the interpretation for standardized variables is that the 'contributions' in the present case are larger than unity.

Two solutions employing test scores in deviate form have been presented. The first, equations (4), gives the principal factors also in deviate form; the second, equations (9), gives the principal factors in standard form. Whether equations (4) or equations (9) are employed it is obvious from the above discussion that the factoring of variances and covariances greatly 'exaggerates' the relative importance of factors, and this may confuse the student who has been used to thinking of all tests as being taken in standard form and therefore having equal unit variances. It is obviously not a question of right or wrong, but merely a choice of metric. The psychologist will probably prefer the analysis of correlations and unit test variances because he can make his tests about equally long and thus may consider them all equally important in the subsequent factor analysis.

We next turn to a point of interpretation which may be made for both the principal-axis and the centroid solutions. This point has to do with the meaning of the second factor, x_2. We shall discuss the issue briefly by referring to our present analysis of variances, but the same issue arises when unit variances and correlations are analyzed by the usual methods. Equations (4)′ will suffice for an illustration. The factor x_1 (here, p_A) is common to both x and y and has positive loadings which seem to be readily interpretable. The factor x_2, however, has a negative loading in the first equation and a positive loading in the second. (These can be reversed by changing the direction of Ox_2.) This means that an increase in score x is accompanied by a decrease in the factor x_2, and vice versa. This may also be readily seen by looking at Fig. 1. As we increase x from the origin (right), the value of x_2 (fourth quadrant) gets larger negatively, that is, farther from its mean or smaller; as we

decrease x from the origin (left), the value of x_2 (second quadrant) gets larger. There is no way of getting around this kind of 'interference' in the principal and centroid solutions. If we change the direction of Ox_2 we merely pass the 'interference' to y and x_2 instead of to x and x_2. In the present analysis there is one positive factor, x_1, and one 'interference' factor, x_2, but with more tests there are generally n-1 such 'interference' factors, where n is the number of tests. The psychologist's great difficulty is in attaching meaning to n-1 such 'interfering' factors. For a problem involving eight physical variables the writer has been able to attach meaning to the second 'interfering' factor, but when he attempted this for some psychological tests, a good deal of mental anguish and confusion arose. I believe a similar confession could be wrung from other psychologists and factor analysts, but this does not mean that I do not approve of 'interference' factors. I approve of them chiefly because of their mathematical elegance, and I simply wait patiently for minds more elastic than my own, and quite a few others, to interpret them so that they may be readily understood.

Although the principal-axis method has long since been derived, it is usually in a form so mathematically forbidding that the student of elementary statistics cannot easily follow it. We shall therefore give a simple derivation in the next section in the ordinary geometry of point representation. In Sections 4 and 5 the principal-axis and centroid solutions will be derived in vector representation introducing some simple new formulas that should clarify the interpretation of both solutions.

3. DERIVATION OF THE PRINCIPAL AXES IN POINT REPRESENTATION

By point representation we mean the ordinary form illustrated by Fig. 2. The twelve pairs of scores in Table 3 have been plotted in this figure as the usual 'scatter diagram.' The ellipse is a contour ellipse whose general equation is

$$\frac{x^2}{v_x} + \frac{y^2}{v_y} - \frac{2p_{xy}}{v_x v_y} = C. \tag{11}$$

The particular ellipse plotted here is one for which $C = 2.5$ is selected arbitrarily so that the ellipse will encompass the points. As the reader probably knows, the principal axes are the axes of the ellipse (11). From the point of view of geometric

simplicity the principal-axis solution is unsurpassed. If Fig. 2 is held so that the Ox axis is horizontal, there is graphic evidence of correlation when the data are thus referred to the Ox and Oy axes. If the paper is rotated through an angle θ so that the

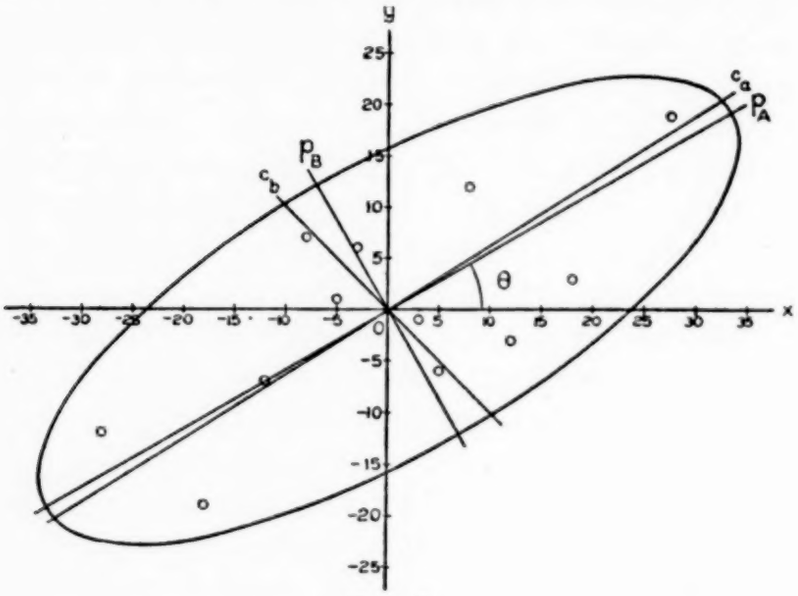

FIG. 2.

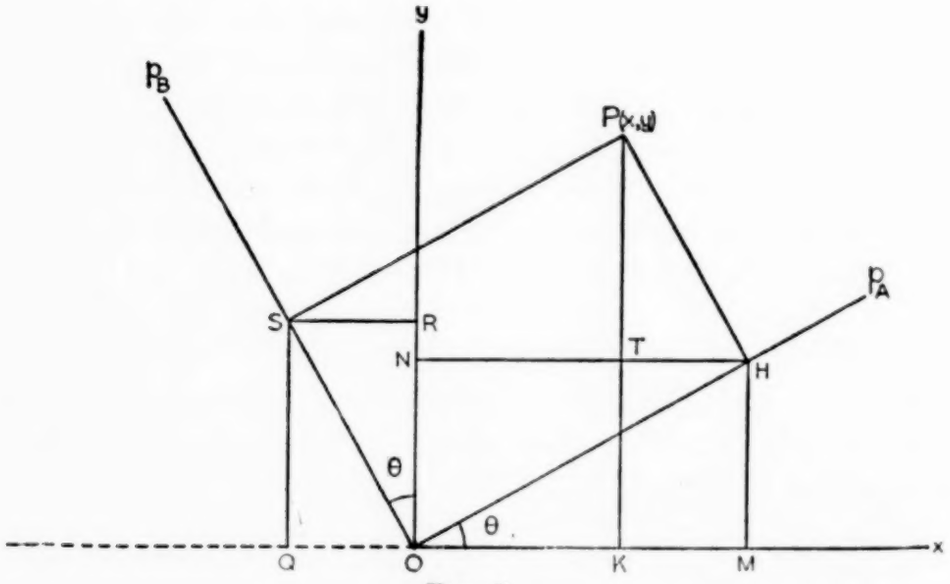

FIG. 3.

Op_A axis is now horizontal, no visual evidence of correlation is apparent. A crude factor analysis has thus been made by merely rotating a piece of paper. Mathematically, we rotate the Ox and Oy axes to new positions, Op_A and Op_B, through an angle θ such that the correlation with respect to the new axes is exactly

TABLE 3.—ILLUSTRATIVE DATA

Case Number	Scores in Deviate Form					
	Original		Principal		Centroid	
	x	y	p_A	p_B	c_a	c_b
1....	28	19	33.75	2.47	34.11	1.32
2....	18	3	17.09	− 6.39	15.24	− 7.17
3....	12	− 3	8.90	− 8.59	6.53	− 9.14
4....	8	12	12.92	6.40	14.52	6.10
5....	5	− 6	1.33	− 7.70	− .73	− 7.94
6....	3	− 1	2.10	− 2.36	1.45	− 2.50
7....	− 3	6	.40	6.70	2.18	6.86
8....	− 5	1	− 3.83	3.36	− 2.90	3.59
9....	− 8	7	− 3.43	10.06	− .73	10.44
10....	−12	− 7	−13.89	− .07	−13.79	.43
11....	−18	−19	−25.08	− 7.47	−26.85	− 6.76
12....	−28	−12	−30.25	3.59	−29.03	4.77

zero. The angle which will accomplish this is given by the formula in Yule[1] as

$$\tan 2\theta = \frac{2p_{xy}}{v_x - v_y}, \tag{12}$$

where $v_x > v_y$. Rather than accept this formula here, we shall derive it by maximizing or minimizing certain variances as described below.

The principal axes are illustrated in Fig. 3. The line p_B is perpendicular to p_A. The point $P(x,y)$ in this figure is to be considered as any point in the scatter diagram, such as one of the twelve points in Fig. 2. We seek a position of the axis Op_A which will fit the scatter diagram of points as well as possible.

[1] G. Udny Yule, *Introduction to the Theory of Statistics*, p. 321. London: Charles Griffin and Co., 1919.

Such a fit may be obtained by minimizing the sums of squares of the distances $OS = p_B$, or by maximizing the sums of squares of the perpendicular projections of $P(x,y)$ on the Op_A axis; that is, $OH = p_A$. We shall obtain the same result by minimizing Σp_B^2, or $\Sigma p_B^2/N = v_B$, as by maximizing $\Sigma p_A^2/N = v_A$, so we shall give the proof only for the former (the reader should check the alternate proof).

The second of equations (6) is

$$v_B = v_x \sin^2 \theta - 2p_{xy} \sin \theta \cos \theta + v_y \cos^2 \theta. \tag{13}$$

The first and second derivatives of v_B with respect to θ are

$$\frac{dv_B}{d\theta} = 2v_x \sin \theta \cos \theta - 2p_{xy} (\cos^2 \theta - \sin^2 \theta) - 2v_y \sin \theta \cos \theta$$
$$= v_x \sin 2\theta - 2p_{xy} \cos 2\theta - v_y \sin 2\theta, \tag{14}$$

and

$$\frac{d^2v_B}{d\theta^2} = 2v_x \cos 2\theta + 4p_{xy} \sin 2\theta - 2v_y \cos 2\theta$$
$$= 2(v_x - v_y) \cos 2\theta + 4p_{xy} \sin 2\theta. \tag{15}$$

This last expression is positive since $v_x > v_y$ and $\theta < 45°$, so that we shall obtain a minimum value for v_B by setting the right member of equation (14) equal to zero, or

$$\tan 2\theta = \frac{2p_{xy}}{v_x - v_y}, \tag{16}$$

which is the same as equation (12).

The above derivation gives a line known in statistics as the 'mutual regression' line. Instead of taking the 'residuals' parallel to the Ox or the Oy axis as by ordinary methods, such residuals have been taken perpendicular to the required line Op_A. This gives a better fit, or smaller error of estimate. The fact that the principal axis thus fits the data better than does any other method is to the writer its chief claim to statistical distinction.

For the present data, $\tan 2\theta = 216/(225 - 100) = 216/125 = 1.728$. From trigonometric tables 2θ is approximately 59°56′, and the required angle θ is 29°58′, whose sine and cosine are .4995 and .8663, respectively, and $\tan \theta = .5766$. Equations (4) and (5) may then be written in the forms (4)′ and (5)′ presented

earlier. The equation of the line Op_A with respect to the x, y axes is $y = x \tan \theta$, or

$$y = .5766x, \tag{17}$$

and the equation of the line Op_B is

$$y = -\frac{1}{.5766}x = -1.734x. \tag{18}$$

4. DERIVATION OF THE PRINCIPAL-AXIS SOLUTION IN VECTOR REPRESENTATION

In leading up to the present derivation we shall attempt to link it with the preceding one by reconsidering the problem of maximizing the variance v_1 (or minimizing the variance v_2). From Fig. 1 equations (3) may be written in the form

$$\begin{aligned} x &= x_1 \cos \theta_1 - x_2 \sin \theta_2 = OM - OQ \\ y &= x_1 \sin \theta_1 + x_2 \cos \theta_2 = ON + OR. \end{aligned} \tag{3}$$

We may also write general equations similar to (9) of the form,

$$\begin{aligned} x &= (\sigma_1 \cos \theta_1)z_1 - (\sigma_2 \sin \theta_2)z_2 = OM_1 z_1 - OQ_1 z_2 \\ y &= (\sigma_1 \sin \theta_1)z_1 + (\sigma_2 \cos \theta_2)z_2 = ON_1 z_1 + OR_1 z_2, \end{aligned} \tag{19}$$

by substituting $x_1 = \sigma_1 z_1$ and $x_2 = \sigma_2 z_2$ in equations (3), z_1 and z_2 being in standard form. The magnitudes OM, ON, OQ, and OR in equations (3) are variables depending upon the position of $P(x,y)$ in Fig. 1, but OM_1, ON_1, OQ_1, and OR_1 are the magnitudes of these variables when $x_1 = \sigma_1$ and $x_2 = \sigma_2$. It is obvious from the figure that $OM_1^2 + ON_1^2 = v_1 \cos^2 \theta_1 + v_1 \sin^2 \theta_1 = v_1$, which is in effect an identity. By a simple transformation, however, we may express OM_1, ON_1, etc., as functions of the standard deviations of x and y and of new angles α and β, to be presently defined. It will therefore be possible to maximize v_1 in a manner alternate to that used in the preceding section. It will also be possible to derive the centroid solution in terms of OM_1 and ON_1, making the logic of these two factor methods more directly comparable than by the usual procedures.

Returning to equations (19) we shall make the substitution $x = z_x\sigma_x$, $y = z_y\sigma_y$, giving

$$z_x = \left(\frac{\sigma_1 \cos \theta_1}{\sigma_x}\right) z_1 - \left(\frac{\sigma_2 \sin \theta_2}{\sigma_x}\right) z_2 = \left(\frac{OM_1}{\sigma_x}\right) z_1 - \left(\frac{OQ_1}{\sigma_x}\right) z_2$$
$$z_y = \left(\frac{\sigma_1 \sin \theta_1}{\sigma_y}\right) z_1 + \left(\frac{\sigma_2 \cos \theta_2}{\sigma_y}\right) z_2 = \left(\frac{ON_1}{\sigma_y}\right) z_1 + \left(\frac{OR_1}{\sigma_y}\right) z_2, \quad (20)$$

where all variables are now in standard form. Since the variances of z_x and z_y are unity, the sums of the squares of the coefficients of z_1 and z_2 are equal to one. They may therefore be regarded as sines and cosines of certain angles, which will be defined by the following equations:

$$\begin{aligned} z_x &= z_1 \cos \alpha - z_2 \sin \alpha \\ z_y &= z_1 \cos \beta + z_2 \sin \beta. \end{aligned} \quad (21)$$

Note that $p_{z_x z_y} = \cos \alpha \cos \beta - \sin \alpha \sin \beta = \cos (\alpha + \beta) = \cos \phi$ (later expressed as r_{xy}). Making the substitutions $z_x = x/\sigma_x$, $z_y = y/\sigma_y$ in these equations gives

$$\begin{aligned} x &= (\sigma_x \cos \alpha) z_1 - (\sigma_x \sin \beta) z_2 \\ y &= (\sigma_y \cos \beta) z_1 + (\sigma_y \sin \beta) z_2. \end{aligned} \quad (22)$$

Upon comparing equations (19) and (22) it is apparent that

$$\begin{aligned} OM_1 &= \sigma_x \cos \alpha \ (= \sigma_1 \cos \theta_1) \\ OQ_1 &= \sigma_x \sin \alpha \ (= \sigma_2 \sin \theta_2) \\ ON_1 &= \sigma_y \cos \beta \ (= \sigma_1 \sin \theta_1) \\ OR_1 &= \sigma_y \sin \beta \ (= \sigma_2 \cos \theta_2). \end{aligned} \quad (23)$$

These are the desired expressions combinations of which we shall maximize or minimize to obtain the principal-axis and centroid solutions.

Turning first to the principal axes we have already noted that we seek to maximize v_A (or minimize v_B). From Fig. 3 and the expressions (23) the function to be maximized is

$$v_A = OM_1^2 + ON_1^2 = v_x \cos^2 \alpha + v_y \cos^2 \beta, \quad (24)$$

or the function to be minimized is

$$v_B = OQ_1^2 + OR_1^2 = v_x \sin^2 \alpha + v_y \sin^2 \beta.$$

In point representation the quantities OM_1^2 and ON_1^2 may be interpreted as the contributions of v_x and v_y to v_A or as the parts of v_x and v_y 'attributable' to v_A. Numerical illustrations of these magnitudes are given by equations (9)′ and (10), where

$OM_1 = 14.68$, $ON_1 = 8.47$, $OM_1^2 = 215.5$, $ON_1^2 = 71.7$, and $v_A = OM_1 + ON_1 = 287.2$. What we are about to do then is to rederive the above solutions, but in a quite different way from that of Section 3.

In the case of the principal-axis solution it has already been noted that $\theta_1 = \theta_2 = \theta$. A simple check formula relating the two solutions may then be obtained from equations (23), as follows:

$$\tan^2 \theta = \frac{ON_1}{OM_1} \times \frac{OQ_1}{OR_1} = \frac{\sigma_y \cos \beta}{\sigma_x \cos \alpha} \times \frac{\sigma_x \sin \alpha}{\sigma_y \sin \beta} = \frac{\tan \alpha}{\tan \beta}. \quad (25)$$

We shall now derive the principal-axis solution very simply in vector representation. In this form of representation of data (Fig. 4) a test or variable is indicated by a vector such as OP_x.

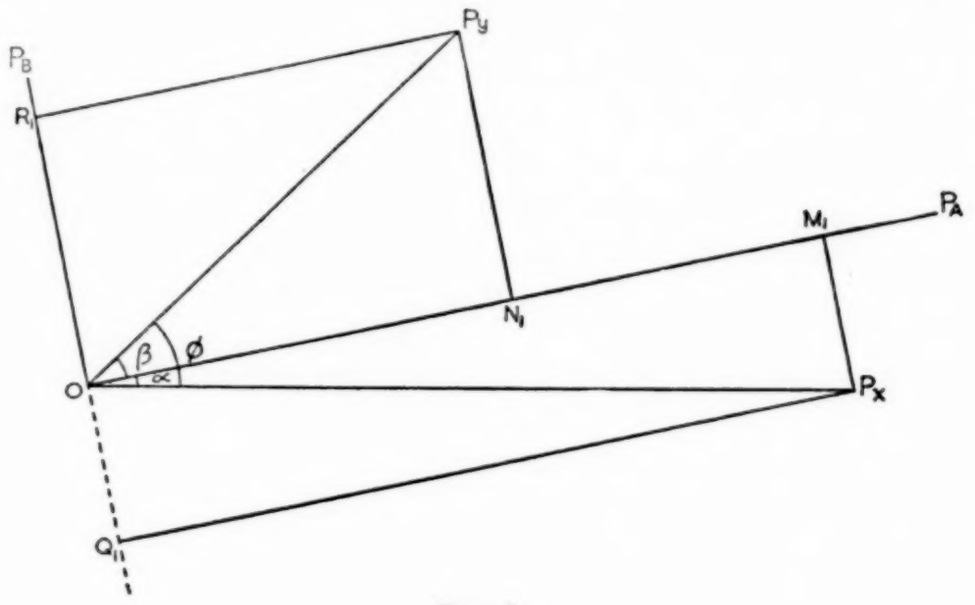

FIG. 4.

The capital letters in this figure are used to denote points. The lengths of such vectors are proportional to the standard deviations of the tests and will here be taken equal to such values for convenience. A line segment OP_x equal to $\sigma_x = 15$ is therefore laid off from the origin O in any direction. Next, a line segment OP_y equal in length to $\sigma_y = 10$ is laid off from the origin and making an angle ϕ with OP_x. The cosine of the angle between any two such vectors is equal to the correlation between the tests x and y. Thus when the angle ϕ is very small, the correlation is very large, approaching 1.00 as ϕ approaches zero. When ϕ

is 90°, the correlation is zero. The present correlation between x and y is .72, so in order to draw our diagram to scale we merely find the angle whose cosine is .72. This angle is approximately 43°57′.

The great simplicity and clarity of the vector form of representation becomes apparent when it is seen that only two coördinates need be thought of here instead of N (12 in our present example). We also have visual evidence of correlation in terms of angular separation of vectors. The magnitudes of variables and factors may also be compared graphically as will presently be illustrated, and both of our factor analyses may be exhibited and easily compared by one simple figure (Fig. 6).

In Fig. 4, M_1 is at the foot of a perpendicular to OP_A from the point P_x, and P_yN_1 is also perpendicular to OP_A. The angle α is that between OP_x and OP_a, while angle β is that between OP_y and OP_A. The sum of these two angles is denoted by ϕ, and cosine ϕ is equal to the correlation $r_{xy} = .72$, as already noted.

The principal axis OP_A might be defined as that vector for which $v_A = OM_1^2 + ON_1^2$ is a maximum or for which $v_B = OQ_1^2 + OR_1^2$ is a minimum. We shall use the former alternate definition in the present proof. We are thus seeking to maximize the contributions of v_x and v_y to the variance v_A, or we may express it as maximizing the contributions of v_A to the variances v_x and v_y as is apparent from equations (10). Employing the first of equations (24) and expressing the angle β in terms of the angles ϕ and α gives

$$\begin{aligned} v_A &= v_x \cos^2 \alpha + v_y \cos^2 \beta \\ &= v_x \cos^2 \alpha + v_y \cos^2 (\phi - \alpha) \\ &= v_x \cos^2 \alpha + v_y(\cos^2 \phi \cos^2 \alpha + \sin^2 \phi \sin^2 \alpha \\ &\qquad + 2 \sin \phi \cos \phi \sin \alpha \cos \alpha) \\ &= v_x \cos^2 \alpha + v_y(\cos^2 \phi \cos^2 \alpha + \sin^2 \phi \sin^2 \alpha \\ &\qquad + \sin 2\phi \frac{\sin 2\alpha}{2}. \end{aligned} \tag{26}$$

The first and second derivatives of v_A with respect to α then become

$$\begin{aligned} \frac{dv_A}{d\alpha} &= -v_x \sin 2\alpha + v_y(-\sin 2\alpha \cos 2\phi + \cos 2\alpha \sin 2\phi) \\ &= -\sin 2\alpha(v_x + v_y \cos 2\phi) + v_y \cos 2\alpha \sin 2\phi, \end{aligned} \tag{27}$$

and

$$\frac{d^2v_A}{d\alpha^2} = -2\cos 2\alpha(v_x + v_y \cos 2\phi) - 2v_y \sin 2\alpha \sin 2\phi. \quad (28)$$

The value of (28) is negative, so that v_A will have a maximum when the right member of equation (27) is set equal to zero, giving the desired equation for obtaining α:

$$\tan 2\alpha = \frac{v_y \sin 2\phi}{v_x + y_y \cos 2\phi} = \frac{\sin 2\phi}{t^2 + \cos 2\phi}, \quad (29)$$

where $t^2 = v_x/v_y$. From the data $v_x = 225$, $v_y = 100$, $\phi = 43°57'$ we find $2\phi = 87°54'$, $\sin 2\phi = .9993$, $\cos 2\phi = .0366$. Substituting the required values in equation (29) then gives

$$\tan 2\alpha = \frac{.9993}{2.25 + .0366} = \frac{.9993}{2.2866} = .4370. \quad (29)'$$

Therefore, 2α is approximately $23°36'$, and $\alpha = 11°48'$. The angle $\beta = \phi - \alpha = 43°57' - 11°48' = 32°9'$. We may then obtain $\sin \alpha = .2045$, $\cos \alpha = .9789$, $\sin \beta = .5321$, and $\cos \beta = .8467$. From equations (23) we then obtain

$$\begin{aligned} OM_1 &= 15 \times .9789 = 14.68 \quad & OQ_1 &= 15 \times .2045 = 3.07 \\ ON_1 &= 10 \times .8467 = 8.47 \quad & OR_1 &= 10 \times .5321 = 5.32. \end{aligned} \quad (23)'$$

The check formula (25) gives

$$\begin{aligned} \tan^2 (29°58') &= \frac{\tan (11°48')}{\tan (32°9')} \\ (.5766)^2 &= \frac{.2089}{.6285} \\ .3325 &\doteq .3324. \end{aligned} \quad (25)'$$

The final pattern is then again obtained in the form (9)′, and the remainder of the solution may be expressed as in equations (10) and Table 2, which have already been interpreted.

5. DERIVATION OF THE CENTROID SOLUTION IN VECTOR REPRESENTATION

A diagram similar to Fig. 4 is shown in Fig. 5 for the centroid solution, where the projections and angles are indicated by primes to distinguish them from the principal axis values. The equations of the centroid axes with respect to the Ox and Oy axes will be given later.

The centroid solution may now be defined as one for which the average (or total) values of the coördinates of P_x and P_y with respect to OP_a and OP_b are maximal; that is, the axis OP_a is one for which

$$W_1 + OM_1' + ON_1' \text{ is a maximum,} \tag{30}$$

while the OP_b axis is one for which

$$W_2 = -OQ_1' + OR_1' \text{ is a maximum.} \tag{31}$$

If these coördinates are interpreted in the usual manner as 'loadings,' it is apparent that in the centroid method we seek axes such that the *average* 'loadings' of the tests (here, only two)

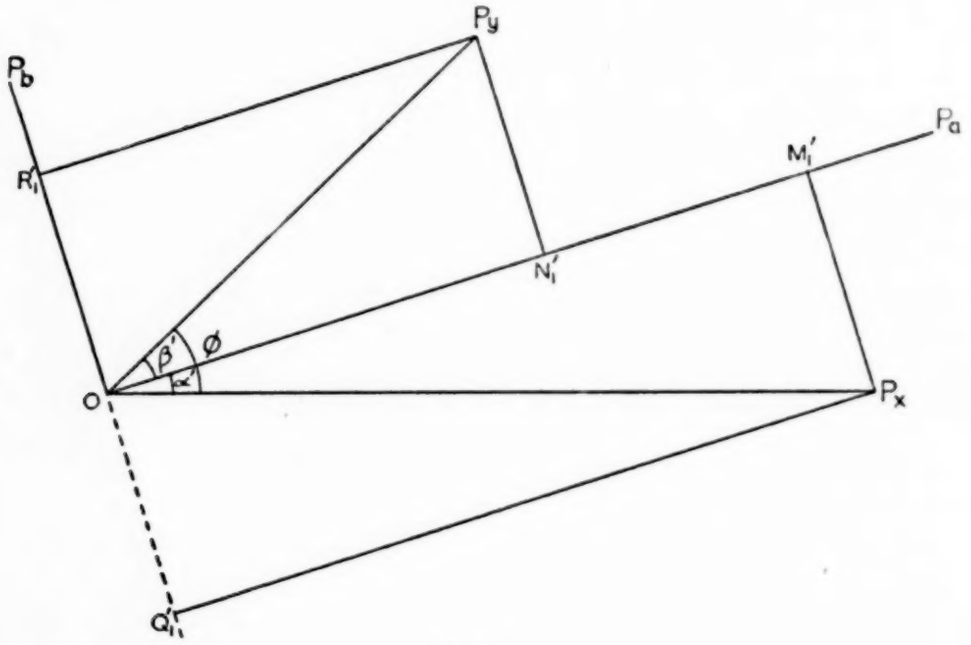

FIG. 5.

are maximized. In the case of the principal factors, however, we obtain axes such that the *sum of the squares* of the 'loadings' are maximized. This is the essential difference between the two methods from the viewpoint of vector representation. The derivation for the centroid method is very simple. We shall give only the OP_a axis from equation (30), but the reader should verify the derivation of the OP_b axis from equation (31). If

$$\begin{aligned} W_1 = OM_1' + ON_1' &= \sigma_x \cos \alpha' + \sigma_y \cos \beta' \\ &= \sigma_x \cos \alpha' + \sigma_y(\cos \phi \cos \alpha' + \sin \phi \sin \alpha'), \end{aligned} \tag{32}$$

the first and second derivatives of W_1 with respect to α' may be written,

$$\frac{dW_1}{d\alpha'} = -\sigma_x \sin \alpha' - \sigma_y \cos \phi \sin \alpha' + \sigma_y \sin \phi \cos \alpha' \quad (33)$$

$$\frac{d^2W_1}{d\alpha'^2} = -\sigma_x \cos \alpha' - \sigma_y \cos \phi \cos \alpha' - \sigma_y \sin \phi \sin \alpha'. \quad (34)$$

Since the second derivative is negative, W_1 will have a maximum when the right member of equation (33) is set equal to zero, giving the simple formula,

$$\tan \alpha' = \frac{\sigma_y \sin \phi}{\sigma_x + \sigma_y \cos \phi} = \frac{\sin \phi}{t + \cos \phi}, \quad (35)$$

where $t = \sigma_x/\sigma_y$. For the present data,

$$\tan \alpha' = \frac{\sin 43°57'}{1.5 + \cos 43°57'} = \frac{.6940}{1.5 + .72} = .3126, \quad (35)'$$

and $\alpha' = 17°22'$. The angle $\beta' = \phi - \alpha' = 43°57' - 17°22' = 26°35'$.

From equations (23) we then find

$$\begin{aligned} OM_1' &= 15 \times .9544 = 14.32 \quad OQ_1' = 15 \times .2985 = 4.48 \\ ON_1' &= 10 \times .8943 = 8.94 \quad OR_1' = 10 \times .4475 = 4.48, \end{aligned} \quad (23)''$$

and the centroid factor pattern will then be

$$\begin{aligned} x &= 14.32z_a - 4.48z_b \\ y &= 8.94z_a + 4.48z_b. \end{aligned} \quad (36)$$

Equations of the form (10) will then furnish the variances of variables and factors in the form

$$\begin{aligned} v_x &= (14.32)^2 + (4.48)^2 = 205.06 + 20.07 = 225.13 \\ v_y &= (8.94)^2 + (4.48)^2 = 79.92 + 20.07 = 100.01 \end{aligned} \quad (37)$$

$$\begin{aligned} v_x + v_y &= 284.98 + 40.14 = 325.12 \\ &= v_a + v_b. \end{aligned}$$

On comparing the variances in these equations with those shown in equations (10) it is apparent that the variance v_a

(285.0) is slightly less than the variance v_A (287.2), while the variances v_b and v_B are of course in reverse order of magnitude.

The equations of the centroid axis with respect to the Ox and Oy axes are obtained from equations (36), as follows:

$$y = .6243x \quad \text{(equation of } OP_a \text{ axis)} \tag{38}$$
$$y = -x \quad \text{(equation of } OP_b \text{ axis),} \tag{39}$$

and these lines are plotted in Fig. 2.

In Fig. 6 the solutions presented graphically in Figs. 4 and 5 have been combined to give a simple geometric interpretation of the two. The relative importance of the factors is indicated by the lengths of the vectors (these being proportional to the standard deviations of the variables they represent). Intercorrelations of variables are easily interpreted visually by the

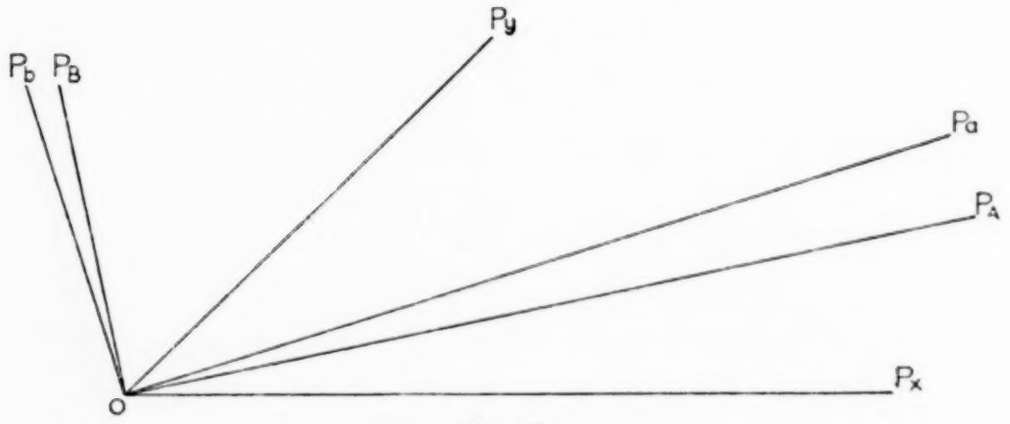

FIG. 6.

angular separation of the vectors. Such a graph gives the clearest graphical representation of the relationships between the principal-axis and centroid solutions here studied.

It is of some interest to note the relationship between methods by determining the relationship between tan 2α and tan α'. From formulas (29) and (35) the following formula may be derived by simple trigonometry:

$$\frac{\tan 2\alpha}{\tan \alpha'} = \frac{2 \cos \phi (t + \cos \phi)}{t^2 + \cos 2\phi} = 1.398. \tag{40}$$

We may therefore find the principal-axis solution from the centroid, or vice versa, but the direct use of (29) or (35) would be simpler.

For $t = 1$, formula (29) reduces to

$$\tan 2\alpha = \frac{\sin 2\phi}{1 + \cos 2\phi} = \frac{2 \sin \phi \cos \phi}{1 + 2 \cos^2 \phi - 1} = \tan \phi, \quad (41)$$

whence $2\alpha = \phi$.

For $t = 1$ and $2\alpha = \phi$, formula (35) then becomes

$$\tan \alpha' = \frac{\sin \phi}{1 + \cos \phi} = \frac{\sin 2\alpha}{1 + \cos 2\alpha} = \frac{2 \sin \alpha \cos \alpha}{1 + 2 \cos^2 \alpha - 1} = \tan \alpha. \quad (42)$$

Therefore $\alpha = \alpha'$. The principal-axis and centroid solutions are thus identical when $t = 1$, the angle θ in point representation being 45° from formula (12).

The question might be asked: "If you are going to factor variances, which method would you use?" The writer's answer would be to choose the principal axes because they fit the data better than do the centroids. The computational labor for the principal axes, however, is generally much greater than for the centroid solution when several variables are factored, and for this reason the centroid method might be preferable for some problems where it is necessary to sacrifice some mathematical nicety.

As a final point we may deal briefly with the vexing question of the 'interference' of the x_2 factor with x or with y. We cannot 'get rid of this interference' if we use only two factors, but we can do so very readily if we add a third. If the three factors in standard form are denoted by F_1, F_2, and F_3, by inspection[1] we may obtain the solution,

$$\begin{aligned} x &= 12F_1 + 9F_2 \\ y &= 9F_1 \qquad + 10\sqrt{.19}\,F_3. \end{aligned} \quad (43)$$

This solution fits the data perfectly, since $p_{xy} = 108$, $v_x = 225$, and $v_y = 100$. It has the advantage of getting rid of the 'interference' completely by considering F_2 and F_3 as unique factors with positive loadings. Of course, there are no fixed rules for

[1] Actually we first wrote

$$\begin{aligned} z_x &= .8F_1 + .6F_2 \\ z_y &= .9F_1 \qquad + \sqrt{.19}\,F_3. \end{aligned}$$

by inspection, and then substituted $z_x = x/15$ and $z_y = y/10$ to obtain (43)

'factoring a doublet' in this manner, and we can write down at once another equally good solution:

$$\begin{aligned} x &= 13.5F_1 + 15\sqrt{.19}\,F_2 \\ y &= \ 8\ F_1 \qquad\qquad + 6F_3, \end{aligned} \tag{44}$$

and an infinitude of others. With more variables, however, a determinate form of solution (Spearman or bi-factor) may be set up and unique solutions obtained free from 'interference' by the addition of an extra factor.

A CASE STUDY OF THE SPLIT-HALF RELIABILITY COEFFICIENT

LEE J. CRONBACH

University of Chicago

Since any test may be split in a large number of ways, the split-half method of estimating test reliability fails to give a unique result. Conventionally, an odd-versus-even-items split is used, which eliminates any act of judgment and leaves the result entirely in the hands of chance. A contrasting procedure referred to as the parallel-split technique has been described by the writer.[3] The tester deliberately divides the items between two half-tests, as much alike as possible in difficulty, variance, and variables measured. The cases on which the split is based are not used in determining the reliability coefficient.

Guttman[4] has recently contributed two ideas to the split-half approach. He offers the formula

$$r_{11} = 2\left(1 - \frac{\sigma_a^2 + \sigma_b^2}{\sigma_t^2}\right)$$

to replace the correlation of half-scores a and b and the Spearman-Brown formula. This formula does not require that the half-tests have equal variance. Guttman then advances the principle that the highest split-half coefficient obtainable (assuming an indefinitely large population) is the best, in the sense of approaching the 'true reliability.' An evaluation of this theory is reserved for a subsequent paper.

It is most important to know more about the variation to be expected among the $\frac{2n!}{2(n!)^2}$ different coefficients obtainable by splitting a test of 2n items. In the absence of a mathematical description of the distribution of these coefficients, empirical results are called for. To date, only Brownwell's results[2] are available. He studied five classroom tests of varying length, obtaining two or three coefficients by random splitting, and about nine coefficients from arbitrary splits of each test. These coefficients showed a wide range. Since most of the splits he made would not ordinarily be used to determine reliability, the extent of variation to be expected in normal practice is uncertain.

PROCEDURE

For definitive conclusions, a variety of tests will have to be studied. The present paper deals with a single test, obtaining a large number of split-half coefficients by the random and parallel-split methods. The test studied was adapted from Booker's unpublished *Test of Silent Reading Vocabulary*, Form 1.[1] The first thirty-eight items of his test were administered to four classes of college psychology students. In each item, the student chooses the best of five suggested synonyms for a moderately difficult word presented in context. Items are arranged alphabetically. Students had ample time to complete the test, and were directed to attempt every item, even when in doubt. The number right was taken as the score.

One hundred papers were used for determining reliability coefficients. Responses were punched on cards, and a table was prepared showing the number passing each item and each pair of items. These data were punched into another set of cards, and by a complex use of the tabulator, sums needed to determine the standard deviation of scores on any combination of items were obtained. This method required about twenty minutes to determine the sums for any split, and limited the test to thirty-eight items. The test-scoring machine would probably have been more adaptable to this problem, had it been available.

RESULTS

The mean score on the test was 19.20; sigma was 6.17. The percentage passing individual items ranged from ninety-nine to eleven. The median item difficulty (all medians were determined from smoothed cumulative frequency curves) was 51; the quartiles were 32 and 67. Table 1 presents relevant data for several arbitrary splits and all random splits. In listing the results, the half of the test having the greater mean was taken as half *a*. All reliability estimates are based on Guttman's formula, which tends to yield results slightly lower than the Spearman-Brown method.

Split 1 is the conventional odd-even split. In Split 2, the first nineteen items are compared with the last nineteen. Splits 3 and 4 are deliberately poor splits, the halves being made dissimilar. The splitting was determined on the basis of an item

analysis of forty-six papers not used in the reliability calculation. In Split 3, the nineteen easiest items were placed in one half. In Split 4, the items having the greatest variance were placed in one half.

The test was split at random thirty separate times, using a table of random numbers. The coefficients, which are presumably representative of the whole population of splits obtainable in practice, are presented in Table 1 in the order in which the splits were made, numbered from 1R to 30R.

Parallel splits were based on the item analysis of forty-six papers not in the main sample. An attempt was made to pair items in the two halves on difficulty and variance. Little attention was paid to content because the items, representative of general vocabulary, are fairly homogeneous in content. Fourteen parallel splits were made, and the resulting data are given in Table 2 in order in which the splits were made (1P to 14P). The splits first made are in general those which seemed most natural; as more splits were made it became increasingly difficult to keep the halves similar, and the splits significantly different from each other. More than fourteen parallel splits can be made for such a test, but these fourteen should be representative of the ones most likely to be made.

For comparative purposes, the reliability was estimated by the Kuder-Richardson formula 20. This yielded a value of .820.

It must be remembered that conclusions may be drawn only for this test, or, at best, for tests having similar statistical properties. This is a short test, with a wide range of item difficulties. It is probable that the item matrix approaches rank one, since the items are similar, with no evident group factors.

Coefficients from random splitting range from .766 to .872; the corresponding coefficients of alienation, from .484 and .358. Lucky or unlucky splits can affect drastically the estimate of reliability and any conclusions based upon it. A standard error of measurement based on the poorest split would be half again as large as that based on the highest split coefficient. In this test, the odd-even split happens to be 'unlucky,' yielding a value of .796 compared to a median of .823 for all random splits.

Random splits often yield half-tests which are not comparable in difficulty and variance. Nearly all of the parallel splits yield half tests whose means are within a point of each other ($M_a \leqq$

TABLE 1.—RELIABILITY COEFFICIENTS AND OTHER DATA OBTAINED BY VARIOUS PLANNED AND RANDOM SPLITTINGS

Split Used	M_a*	σ_a	σ_b	σ_a/σ_b	r_{11}	$\sqrt{1-r_{11}}$
1. Odd-even	10.00	3.44	3.33	1.03	.796	.452
2. First-last	10.57	3.30	3.28	1.01	.862	.371
3. Easiest-hardest	12.91	3.28	3.46	.95	.806	.440
4. Greatest variance-least variance	9.67	3.69	3.01	1.23	.809	.437
Random:						
1R	9.89	3.23	3.42	.94	.839	.401
2R	11.10	3.21	3.55	.90	.798	.449
3R	10.34	3.38	3.32	1.02	.824	.420
4R	10.03	3.36	3.38	.99	.809	.437
5R	10.03	3.57	3.12	1.14	.820	.424
6R	10.64	3.28	3.32	.99	.855	.381
7R	9.72	2.94	3.73	.79	.815	.430
8R	10.25	3.11	3.51	.89	.843	.396
9R	10.47	3.49	3.25	1.07	.805	.442
10R	10.36	3.14	3.50	.90	.841	.399
11R	10.12	3.77	2.89	1.30	.813	.432
12R	10.32	3.45	3.23	1.07	.828	.415
13R	9.65	3.55	3.31	1.07	.766	.484
14R	10.47	3.60	3.10	1.16	.815	.430
15R	9.61	3.12	3.62	.86	.799	.448
16R	10.68	3.17	3.56	.89	.805	.442
17R	9.73	3.52	3.16	1.11	.827	.416
18R	10.29	3.36	3.34	1.01	.821	.423
19R	10.06	3.16	3.46	.91	.847	.391
20R	11.06	3.33	3.35	.99	.828	.415
21R	10.16	3.02	3.58	.84	.848	.390
22R	9.67	3.34	3.26	1.02	.858	.377
23R	9.87	2.88	3.72	.77	.840	.400
24R	10.80	3.10	3.62	.86	.807	.439
25R	11.07	3.24	3.53	.92	.795	.453

TABLE 1.—RELIABILITY COEFFICIENTS AND OTHER DATA OBTAINED BY VARIOUS PLANNED AND RANDOM SPLITTINGS—(*Continued*)

Split Used	M_a*	σ_a	σ_b	σ_a/σ_b	r_{11}	$\sqrt{1-r_{11}}$
Random:						
26R	10.13	3.15	3.48	.91	.842	.397
27R	9.89	2.99	3.81	.78	.769	.481
28R	9.79	3.57	3.22	1.11	.787	.462
29R	9.73	3.39	3.16	1.07	.872	.358
30R	10.36	3.29	3.40	.97	.824	.420
Median**	10.14	3.28	3.40	.97	.823	.420
Q_1	9.83	3.14	3.25	.88	.804	.443
Q_3	10.41	3.42	3.53	1.04	.837	.404

* Would be 9.60 if $M_a = M_b$.

** Medians and quartiles determined from smoothed cumulative frequency curves.

10.10), and whose sigmas have a ratio between .90 and 1.10. Of the thirty random splits, only six (1R, 4R, 13R, 19R, 22R, and 29R) are thus similar. The coefficients obtained in these cases are .839, 809, .766, .847, .858, and .872. Four of these are among the highest fourth of the coefficients, but the other two are low.

The Spearman-Brown formula, and the logic of the split-half approach, require half-tests to be comparable. In Table 3, coefficients are grouped according to the similarity of the variances of the half-tests. Splits for which the ratio of half-test sigmas is between .90 and 1.10 are separated from those less comparable. Coefficients based on comparable half-tests tend to be larger than those from non-comparable splits, but there is a wide range of values from the comparable splits.

The parallel-split method is no more successful than the random split in obtaining uniform coefficients. The fourteen coefficients range from .766 to .858, and the inter-quartile range is .042, compared to .033 for the random splits. Three of the splits (6P, 8P, and 12P) were not adequately comparable, probably because of sampling errors in the data on which the split was based. These splits account for two of the lowest parallel-split coefficients. A progressive decline in the comparability of the test halves from the first to the last parallel split is to be noted.

For these data, the parallel split produces halves of equal dif-

TABLE 2.—RELIABILITY COEFFICIENTS AND OTHER DATA OBTAINED FROM PARALLEL SPLITTINGS

Split	M_a*	σ_a	σ_b	σ_a/σ_b	r_{11}	$\sqrt{1-r_{11}}$
1P	9.61	3.22	3.51	.92	.808	.438
2P	9.67	3.40	3.28	1.04	.826	.417
3P	9.73	3.44	3.23	1.07	.832	.410
4P	9.96	3.34	3.26	1.02	.857	.378
5P	9.71	3.45	3.14	1.10	.858	.377
6P	10.13	3.02	3.78	.80	.774	.475
7P	9.79	3.21	3.44	.93	.838	.402
8P	9.83	3.61	3.08	1.17	.815	.430
9P	9.76	3.23	3.46	.93	.824	.420
10P	9.94	3.24	3.41	.95	.839	.401
11P	9.99	3.33	3.44	.97	.799	.448
12P	10.10	3.20	3.64	.88	.766	.484
13P	9.92	3.17	3.44	.92	.851	.386
14P	9.87	3.25	3.51	.93	.797	.451
Median**	9.86	3.27	3.42	.94	.825	.418
Q_1	9.72	3.19	3.28	.91	.798	.449
Q_3	9.98	3.39	3.52	1.04	.840	.400

* Would be 9.60 if $M_a = M_b$.

** Medians and quartiles determined from smoothed cumulative frequency curves.

TABLE 3.—RELIABILITY COEFFICIENTS IN RELATION TO THE EQUIVALENCE OF HALF-TESTS

Split	σ_a/σ_b	n	Statistic	M_a	r_{11}
Random	.90–1.10	17	Median	10.28	.829
			Q_1	9.98	.808
			Q_3	10.56	.844
Random	>.90; <1.10	13	Median	10.01	.815
			Q_1	9.78	.803
			Q_3	10.33	.828
Parallel	.90–1.10	11	Median	9.82	.832
			Q_1	9.69	.810
			Q_3	9.95	.850

ficulty ($M_a \leqq 10.10$) and variance ($.90 < \sigma_a/\sigma_b < 1.10$) eleven times out of fourteen, while halves from random splits meet this test only six times out of thirty. Random halves are comparable in variance alone in seventeen splits out of thirty. When splits yielding comparable half-tests are considered (Table 3) the parallel split shows a very slight advantage over the random splits.

DISCUSSION

One has no basis for judging the 'true' reliability coefficient. If one accepts the split-half procedure as legitimate, it is fair to consider the reliability as greater than .829, the median of the comparable random splits. The split-half method attempts to predict the correlation that would be obtained between two tests as similar to each other as possible. Because of possible sampling errors, one cannot assume that such tests would correlate as well as the two halves in split 29R, which yielded a coefficient of .872. Since several coefficients fall above .850, it seems fair to accept a value in the neighborhood .850 to .870 as the best estimate of reliability from these data.

The results of this study suggest possible procedures for future use of the split-half principle:

1) A single random, or odd-even, split gives results of limited usefulness, since it may be atypical. There is little chance of securing two consecutive atypical coefficients which agree with each other. Every test should be split in at least two ways. If the coefficients agree, and the half-tests are equivalent, the result is probably near the median split-coefficient for that test and sample. If the coefficients do not agree, further splits are required to establish a stable value. There is no practical procedure for determining the highest meaningful coefficient, but it should be recognized that the median split-half coefficient is probably an underestimate of the correlation of truly equivalent tests.

2) Halves obtained in splitting are likely to be non-comparable. Splits must be checked for comparability, and discarded if they fail to yield equivalent half-tests. Writers reporting split-half coefficients should report the means and sigmas of the half-tests.

3) For this test, the parallel-split offers little advantage, the results being similar to those for random, comparable halves. The parallel split does yield a higher proportion of comparable splits, and may save time in practice.

Further studies of this type are required. In particular, information is needed regarding longer tests, tests of low reliability, and tests not measuring a single general factor. The parallel split is likely to be more advantageous for tests of the latter sort than for the test used here.

REFERENCES

1) Booker, Ivan A, "The Measurement and Improvement of Silent Reading among College Freshmen." Unpublished doctor's thesis, University of Chicago.

2) Brownell, W. A. "On the accuracy with which reliability may be measured by correlating test halves." *J. Exper. Edu.*, Vol. I, 1933, pp. 204–215.

3) Cronbach, Lee J. "On estimates of test reliability." *J. Educ. Psychol.*, Vol. XXXIV, 1943, pp. 485–494.

4) Guttman, Louis. "A basis for analyzing test-retest reliability." *Psychometrika*, 1945, Vol. X, pp. 255–282.

THE FACTORIAL COMPOSITION OF TWO TESTS OF COMPREHENSION IN READING

FREDERICK B. DAVIS

Office of the Air Surgeon

Washington, D. C.

In 1942, the writer described two new measures of reading ability, Word Knowledge and Reasoning in Reading.[1] These measures were identified by means of factorial techniques and provided the first practical results of the application of factorial analysis to the measurement of reading ability.

Since that time, additional information has become available regarding the psychological composition of these two important components of comprehension in reading—information that may help us to understand a little more clearly what mental skills are called into play in the process of comprehending. It is appropriate to recall that the writer's original data were based on the scores of 421 college freshmen who were given sufficient time to try every item in Form Q of the "Coöperative Reading Comprehension Tests." Therefore, speed of reading did not influence the scores that were used for analysis and the results must be interpreted with this fact in mind. Most studies of reading ability have been based on scores that represented a mixture of speed of reading and level of comprehension.

In 1943, the writer included in a battery of short tests of reasoning and judgment two tests that played the largest part in determining the Word-Knowledge and Reasoning-in-Reading components reported in 1942. These two tests consisted of ten recognition-vocabulary items and ten inferences-in-reading items. The entire battery of tests was administered with a very generous time limit to eleventh- and twelfth-grade boys in New York City, and the answer sheets of 689 boys who tried every item were used for purposes of analysis. Again, speed of reading had no influence on the scores. A complete report of the analysis of the Judgment and Reasoning Tests has been prepared[2] and

[1] F. B. Davis, "Two New Measures of Reading Ability," *J. Educ. Psychol.*, Vol. 33, No. 5 (May, 1942), pp. 365–372.

[2] *Vid.* F. B. Davis, ed., *The AAF Qualifying Examination.* AAF Aviation Psychology Program Research Reports, No. 6. Washington: Government Printing Office, 1947. Chapter 7.

no attempt will be made to duplicate the content of that report. Only that portion of the data that is relevant to an analysis of comprehension in reading will be discussed here.

Intercorrelations of the fourteen judgment and reasoning tests were computed and a factorial analysis of the resulting matrix was made. The fourteen variables were weighted equally for purposes of analysis, each diagonal entry being unity, and the principal-axis method described by Kelley was employed.[3] Eight independent components were obtained in which individual scores were found to have reliability coefficients significantly greater than zero at the five per cent level, or better. This indicated that the variance of each one of the eight components was significantly different from zero. The variance of one other component was found to be significantly different from zero at very close to the same fiducial level. The application of one of Hoel's tests for minimum rank in component analysis indicates that the chances are about nineteen to one that five components are required to explain the intercorrelations in the matrix and that two or three others may be required. Hoel mentions the fact that the test used tends to err on the conservative side.[4] In any case, a precise comparison of the results of applying the tests for minimum rank proposed by Hoel and by the writer[5] cannot be made on the basis of the data derived from the judgment and reasoning study because the reliability coefficients for the fourteen variables were not based on the scores of the 689 pupils used for factorial purposes, but were calculated from the records of 1,026 aviation students and corrected appropriately for differences in range of talent.

Among the eight statistically significant components were two of which the loadings bore a striking similarity to those of the Word-Knowledge and Reasoning-in-Reading components reported in 1942. Since these two components were found to account for a large proportion of the variance of comprehension in reading, it is interesting to note the product-moment correlation

---

[3] T. L. Kelley, *Essential Traits of Mental Life.* Cambridge: Harvard University Press, 1935.

[4] P. G. Hoel, "A Significance Test for Minimum Rank in Factor Analysis," *Psychometrika*, Vol. 4 (December, 1939), pp. 245–253.

[5] F. B. Davis, "The Reliability of Component Scores," *Psychometrika*, Vol. 10 (March, 1945), pp. 57–60.

coefficients between scores in these two components and scores in each one of the fourteen tests in the judgment-and-reasoning battery. These correlations are shown in Table 1.

TABLE 1.—PRODUCT-MOMENT CORRELATION COEFFICIENTS BETWEEN SCORES IN TWO UNCORRELATED COMPONENTS OF COMPREHENSION IN READING AND SCORES IN FOURTEEN TESTS IN THE JUDGMENT-AND-REASONING BATTERY

Test	Product-moment r with	
	Word-Knowledge Component	Reasoning-in-Reading Component
1. General vocabulary	.63	−.34
2. Judgment	.07	.13
3. Pure judgment	−.09	−.07
4. Mechanical judgment	.05	.12
5. Reasoning judgment	−.21	−.14
6. Logical reasoning	.03	.02
7. Arithmetical reasoning	.11	−.32
8. Practical information	.06	.17
9. Mechanical comprehension	.03	.13
10. Inferences in reading	.35	.63
11. Syllogistic reasoning	.19	.05
12. Mechanical movements	−.40	.10
13. Figure analogies	−.25	−.12
14. Pattern analogies	−.20	.20

The Word-Knowledge component is, of course, positively correlated with the recognition-vocabulary and inferences-in-reading tests. It is negatively related to the predominantly nonverbal reasoning tests and the mechanical-movements test, probably because of the mechanical-interest variance in the latter. By and large, there is little relationship between the Word-Knowledge component and several different types of reasoning and judgment items expressed in verbal form. The Reasoning-in-Reading component is strongly (and positively) correlated

with the inferences-in-reading test. It is negatively related to the recognition-vocabulary test and the arithmetical-reasoning test. With other tests of judgment and reasoning it has only slight relationship.

On the basis of these data, it appears that the type of reasoning ability that is most important in reading is not like that demanded by conventional reasoning tests, such as those composed of practical mathematical problems. These conclusions are very interesting from both theoretical and practical points of view. They indicate that reasoning ability, as we usually think of it, is not a generalized trait that can be applied to the solution of any kind of problem. Instead, it is a composite of rather unrelated abilities to reason in different kinds of mental activities. Thorndike's early belief that reading is a process of reasoning closely similar to that involved in solving arithmetic problems is not entirely confirmed.[6] However, his idea that reading does involve skill in some type of reasoning is definitely supported by the data.

Reading seems to be primarily a process of associating word meanings, but it does require reasoning of a highly specialized type, which can probably be developed only by training and experience in reading. Sophistication in formal logic, skill in solving mathematical problems, and skill in syllogistic reasoning are often developed concurrently with ability to reason in reading, but there is no basis for believing that training in any one of them will, by itself, improve an individual's ability to read with understanding. Learning exercises intended to improve a pupil's ability to reason in reading should apparently consist of questions based on reading matter that will require the pupil to weave together the ideas throughout a section of the text and to make inferences from the ideas expressed.

Another approach to understanding the mental abilities required in reading is to examine the factorial composition of word-knowledge and inferences-in-reading items, in terms of the seven nameable components found to have reliability coefficients significantly greater than zero at the five per cent level. Table 2 shows these data for the recognition-vocabulary items and Table 3 for the inferences-in-reading items.

---

[6] E. L. Thorndike, "Reading As Reasoning: A Study of Mistakes in Paragraph Reading," *J. Educ. Psychol.*, Vol. 8 (June, 1917), pp. 325–332.

TABLE 2.—FACTORIAL COMPOSITION OF RECOGNITION-VOCABULARY ITEMS

Components	Regression Coefficient (Factor Loading)	Percentage of Recognition-Vocabulary Variance
1. Practical reasoning ability.....	.22	5
2. Mechanical vs. nonmechanical background.....	.16	3
3. Verbal facility.....	.64	41
4. Pure judgment.....	−.20	4
6. Perceptual facility.....	.25	6
10. Reasoning in reading.....	−.45	20
13. Computational facility.....	−.41	16
Total.....		95

The three negative regression coefficients in Table 2 indicate that scores on the recognition-vocabulary test are negatively related to scores in the components named Pure Judgment, Reasoning in Reading, and Computational Facility. Hence, about forty per cent of the variance of the recognition-vocabulary items is used in suppression variables to cancel out word-knowledge variance in other tests that serve mainly to determine scores in those three components. By far the largest remaining part of the recognition-vocabulary variance (forty-one per cent) constitutes the main element in determining scores in the Verbal Facility component. Small portions of the recognition-vocabulary variance may be ascribed to practical reasoning ability and to perceptual facility. It is well known that one may answer multiple-choice test items by successively eliminating wrong answers and figuring out which choice the test constructor must have intended to be correct. In straightforward vocabulary items this procedure should be relatively unimportant, and so it turns out to be.

The data in Table 3 indicate that the variance of inferences-in-reading items is largely accounted for by the Reasoning-in-Reading component. Ability to make inferences in reading is positively associated with a nonmechanical background and with verbal facility. Small amounts of practical reasoning and

TABLE 3.—FACTORIAL COMPOSITION OF INFERENCES-IN-READING ITEMS

Component	Regression Coefficient (Factor Loading)	Percentage of Inferences-in-reading Variance
1. Practical reasoning ability.....	.23	5
2. Mechanical vs. nonmechanical background.....	.36	13
3. Verbal facility.....	.33	11
4. Pure judgment.....	.04	0
6. Perceptual facility.....	.21	4
10. Reasoning in reading.....	.72	52
13. Computational facility.....	.14	2
Total.....		87

perceptual facility are also involved in responding to inferences-in-reading items.

The fact that perceptual facility is so minor an influence in determining scores in two important elements of ability to read constitutes evidence that perceptual skills are not important in reading when high-school pupils are allowed to take as much time as they want and can proceed at their own rate. Only if speed of silent reading is an important goal of instruction can attention to the mechanics of perception be justified. From these data one may perhaps justifiably conclude that instruments such as the Metron-O-Scope and the Ophthalmograph can not be expected to have sufficient influence on a pupil's power of comprehension to justify the expense and trouble of using them. Whether diagnosis on the basis of the Ophthalmograph or remedial training by means of the Metron-O-Scope, or such devices as the tachistoscope, will increase a pupil's speed of reading more than mere conscious effort to do so is not even touched upon by the data reported above. If so, justification for improving speed of reading would rest on the demonstration of the social utility of ability to read rapidly. Presumably such demonstration could be made in the case of at least a few vocations.

READING REVERSALS AND LATERAL PREFERENCES IN A GROUP OF INTELLECTUALLY SUBNORMAL BOYS

ALEXANDER MINTZ

Department of Psychology, The City College of New York

Reading disability, i.e. the apparent inability of occasional children to learn to read which could not be explained in terms of their intelligence test performance, has been investigated for a long time. In 1925 Orton[16] formulated a new theory for it in neurological terms. According to this theory, a visual shape leaves two memory traces in the visual association center of the cortex, one in each cerebral hemisphere, the two being mirror images of each other. In normal people one cerebral hemisphere is dominant, and its trace consistently determines behavior. In occasional individuals, neither hemisphere is consistently dominant, and either of the two memory traces may be reproduced on any given occasion, which leads to confusion. Thus, according to Orton, 'b' and 'd' being mirror images of each other have identical pairs of memory traces, but in different cerebral hemispheres. If neither hemisphere is dominant, both letters are apt to revive either of the same two memory images, and the subject has no way of identifying either of the two letters consistently. The resulting confusion in case of pairs of letters or words where a similar relationship exists leads to failure to learn to read. In view of the outstanding rôle of the confusion between shapes and their mirror images, Orton gave reading disability a new name, 'strephosymbolia' (affliction with twisted symbols).

Orton's theory was originally formulated on the basis of an investigation of fifteen cases of reading disability, and a few years later his collaborator, Monroe[14] published certain experimental results which were in agreement with it. The principal findings supporting the theory were: the common occurrence of reversals of the sequence of letters in reading words and of confusions between letters and their mirror images (Orton); the tendency of some poor readers to produce mirror writing (Orton); confused handedness and clumsiness of some subjects (Orton); the excess of reversal errors in the reading performance

of retarded readers as compared to normal beginners with the same degree of reading skill (Monroe); relative superiority of retarded readers over normal readers in mirror reading, mirror writing and the use of the non-preferred hand.

The experiments to be reported in this paper were designed to verify the facts reported by Orton and Monroe and to investigate the validity of the theory by testing its consequences. If both right-left reversals of single letters (subsequently referred to as R-errors) and reversals of the order of letters (O-errors) usually are unavoidable effects of a lack of unilateral cerebral dominance, the proneness towards errors of both these kinds should be intercorrelated to a substantial extent. In addition, there ought to be, if Orton's theory is correct, a statistical association between weak or inconsistent lateral preferences in motor performance and reversal errors. There are good reasons to believe that laterality in motor performance is indicative of unilateral cerebral dominance; e.g. it is well known that aphasia is generally caused by left-sided brain injuries in right-handers, by right-sided injuries in left-handers. Both Orton and Monroe point out that weak lateral motor preferences are common among retarded readers; however, if Orton's theory is correct, weak lateral preferences should be primarily associated with reversals rather than with other types of deficient reading. The problems of this study were:

Can the facts described by Orton and Monroe be duplicated by studying the reading performances of a group of children in an institution for the feebleminded, many of whom are retarded readers even in terms of their subnormal intelligence test performance? Are the R-L reversals and the O reversals intercorrelated? Are the children with weak lateral preferences particularly apt to make R-L and O errors?

The experiments were conducted in Letchworth Village, a New York State institution for the feebleminded, in the early part of 1931, and the results (not completely analyzed) were reported at the meeting of the New York Branch of the American Psychological Association in 1933. They were not in favor of Orton's theory. In this they were not unique, inasmuch as Monroe in a later study[15] and Wolfe[24] also obtained results opposed to Orton's theory.

Nevertheless, a survey of recent literature indicated that

Orton has repeatedly restated his theory, at least as recently as 1943,[17] and that the theory continues to exert an influence. It was thought, therefore, that a publication of the results of these experiments was desirable in spite of their age.

SUBJECTS AND PROCEDURE

The subjects were ninety-five boys, all inmates of Letchworth Village. Their ages ranged from seven years, eleven months to seventeen years, three months. All attended school at the institution, the highest grade being the fourth. Their Stanford Binet (1916 edition) IQ's ranged from 47 to 87. About a third of the boys had IQ's above 70, the conventionally accepted upper limit of mental deficiency; since this upper limit is too high, it appears that the group is more accurately described as intellectually subnormal rather than as feebleminded. The presence of a large number of boys with IQ's above 70 in an institution for the feebleminded was probably due, at least in part, to emotional disturbances causing the boys to test below their true level when brought to Children's Court.

Two reading tests were used; all boys were given Monroe's Iowa Reading Test; somewhat more than half were also given the Gray Oral Reading Test. The two tests were found to correlate to the extent of .83 in our sample. Since the Gray test could not be scored in a substantial number of cases of very poor readers, Monroe's test scores, transmuted into grade equivalents, were used in the subsequent computations; in arriving at the grade equivalents the table provided by Monroe was used.

Monroe's procedure with one modification was followed in determining the numbers of reversal errors: four different types of reversals were tabulated separately, while Monroe distinguished between two types and combined their frequencies in her computations. In addition to the O-errors and the R-errors, two other categories of errors were introduced: confusions between letters differing in orientation in the vertical direction, e.g. b and p (top-bottom reversals, subsequently referred to as T errors) and reversals involving both the right-left and the top-bottom direction, or inversions (I-errors), e.g. confusion between n and u or p and d. No distinction was made, as

later by Wolfe, between complete order reversals (e.g. confusion between saw and was) and partial ones (e.g. confusion between form and from). In some instances words involved combinations of two reversal errors, e.g. squirt read as spent. At times different combinations of reversals could account for the same erroneous reading, e.g. pod read as dog. This could be viewed as an O-reversal and a confusion between p and g, or as an I-error (p and d) and a confusion between d and g. Half errors were scored in such cases. It may be added that the interpretation of confusions between b and d and between b and p as errors based on identity of shapes differing in orientation is not altogether convincing; the sound pairs are also rather similar in these instances.

The sighting preference of the boys was determined by a slight variation of the procedure described by W. R. Miles.[12] Thin cardboard cones, identical in shape and size with Miles' 'V-scope' were prepared. In order to be used, they had to be held between two hands and opened. The subjects were shown how to hold the cone and were told to look through it at objects held by the experimenter and name them. According to Miles, practically all adults, including the highly educated, fail to recognize that the shape of the cone forces them to sight with one eye only; most likely, our intellectually subnormal subjects failed to recognize it too. The sighting eye can be determined at once by observing the subject (see, e.g. the photographs in Miles' paper). The sighting test was repeated six times in succession with most subjects, more than six times with a few. The subjects were classified as sighting with the right or left eye, respectively, if the same eye was used each time, with not more than one exception, as impartially sighting in case of the less uneven distributions of the observed sighting eyes.

The handedness was determined in the case of writing, throwing and batting. The writing of the subjects with both hands was observed. Each subject was asked to show how he bats a baseball (only three subjects did not play baseball). Each subject was asked which hand he used in throwing; in addition, he was asked to throw a ball to the experimenter, first with one hand, then with the other; a full throwing motion was required rather than an underhand toss. The throwing efficiency was roughly described in qualitative terms, e.g. "right-handed

throwing satisfactory, left-handed helpless," "poor thrower with both hands, definitely worse with left hand."

In subsequent discussions, four capital letters are used to designate the results of the four laterality tests—sighting, writing, throwing, batting in this order, R referring to the right side, L to the left, I to impartial preference. RRRL, for example, refers to a right-sighting boy who wrote and threw right-handed, batted left-handed.

RESULTS

1) Numerous errors of the two types described by Orton were observed in the reading of the children. Errors classified as R-reversals or as combinations of R-reversals with other types including such examples as the reading of dig as big, of ball as doll, of red as did, of squirt as spread, pod as good, dug as bug, and so on. Balk read as block, of as for, blind as bill, parlor as player, form as from, tap as pot, nip as imp, and many others were classified as order reversals.

2) The two varieties of reversal errors involving the vertical direction were also found, e.g. pig was sometimes read as big (a p becomes a b if the top and the bottom are reversed), dig as pig (a p is a d turned 180°), pod as bad, pardon as bird, plea as bleed, blond as blowed (presumably the subject mistook the n for a u, and then mispronounced the 'bloud'). The T-errors and I-errors were somewhat less numerous in our material than the O and the R-L errors. Still, considerable numbers of them were found.

3) The verification of Monroe's finding that reversals were more commonly made by retarded readers than by normal beginners in reading presented difficulties. Too many of our poor readers were retarded, too many of the better ones advanced in reading; the correlation between reading grade and reading retardation in our material was −.61. The correlations between reading retardation and the frequencies of the O, R, T and I errors were all positive, 0.56 for O-reversals, 0.26 for R-reversals, 0.42 for T reversals, 0.21 for I-errors; in part, these correlations are presumably due to the fact that these error frequencies and reading retardation are all correlated negatively with the reading grade. The partial correlations with the reading grade held constant, between reading retardation and the error frequencies were:

for O errors: .48 ± .056
for R errors: .02
for T errors: .22 ± .069
for I errors: .00

The partial correlations duplicate Monroe's results. Retarded readers made more reversal errors of two types than could be explained in terms of the correlations between reading grade and reversals and between reading grade and reading retardation. Clearly, the same result would have been obtained if all four reversal frequencies had been combined, as had been done by Monroe. Since Monroe does not present the data pertaining to the frequencies of the various types of reversals, it cannot be determined whether the same reversal types correlate with reading retardation in her material as in ours.

4) Table I presents the intercorrelations between the frequencies of the four kinds of errors previously characterized. The correlations were computed separately for different levels of reading skill, and also for the group as a whole.

TABLE I.—INTERCORRELATIONS BETWEEN FREQUENCIES OF FOUR KINDS OF ERRORS.

Number of Boys	Read. Grade	r_{OR}	r_{OT}	r_{OI}	r_{RT}	r_{RI}	r_{TI}
16	1–1.3	.63	.01	.21	−.20	.20	.44
18	1.4–1.9	−.08	−.02	.10	−.09	.25	−.16
30	2–2.9	.05	.34	.54	.16	.15	.28
22	3 and above	−.06	.36	.39	−.18	−.12	.03
86*	1–or ca. 5.5	.32	.32	.37	.19	.28	.28

* The boys whose reading appeared to be almost pure guessing, or who did not respond at all to many of the words were not forced to complete the test; their records are eliminated from the tables based on error counts.

In the total group, all three correlations involving O-errors are somewhat higher than the other three correlations; r_{OR}, involving the two types of errors emphasized by Orton in his theory, is low in absolute amount and is no higher than either of the other two correlations involving O-errors. The correlations between the two types of errors, the reversals of orientation of single letters are somewhat lower. They form an under-

standable pattern, inasmuch as the errors not involving the same direction (R errors and T errors) are intercorrelated to a lesser extent than the other two pairs, the R errors and I errors (both involving horizontal reversals) and T errors and I errors (both involving reversals in the vertical direction). The two last mentioned pairs of variables are correlated to the same extent, $r_{RI} = r_{TI} = .28$. All of the differences are of doubtful statistical significance at best; the essential finding is the failure of the correlations based on confusions between right and left to stand out in any way among the other correlations.

The correlations within the groups classified according to reading level fluctuated irregularly, as one should expect, in the case of small groups.

Partial correlations, computed between the error frequencies, with the reading grade partialled out, essentially confirmed the above findings. They were:

Between O and R: + .12
Between O and T: − .12
Between O and I: + .28
Between R and T: − .04
Between R and I: + .18
Between T and I: + .15

Only one of the correlations is statistically significant, that between O errors and I errors. In spite of their lack of trustworthiness, it may be noted that the correlations form a plausible pattern: the correlations are positive in the case of pairs of variables, both involving the right-left direction (any combination of O, R and I) or the vertical direction (T and I). The signs are negative in the cases of pairs of variables not involving the same direction, viz. the correlations of T with O and R, respectively. The correlation between R and O again fails to stand out. The correlations of the other variables involving the same directions, either the horizontal or the vertical, are actually a trifle higher.

5) Consistent preference for the right side was found to be somewhat associated with superior reading performance. Table II presents the numbers of RRRR and RRR-cases (i.e. the numbers of pure and probably pure right-handers sighting

with the right eye) and the numbers of all other laterality types at different levels of reading retardation or acceleration, respectively.

TABLE II.—NUMBERS OF RRRR AND RRR-CASES

Reading Retard. or Accel.	Number of RRRR and RRR-	Number of All Others
Retarded 1.5 gr. or more.....	4 (18%)	18 (82%)
Retarded .5–1.4 gr.....	7 (25%)	21 (75%)
Adv. or ret. less than .5 gr.....	10 (33%)	20 (67%)
Adv. .5–1.4 gr.....	5 (50%)	5 (50%)
Adv. 1.5 gr. or more.....	2 (40%)	3 (60%)

The differences between the numbers of right-sighting right-handers and chance expectation are not large, the chi square test indicating that differences of this (or greater) magnitude, occurring in a random order, have a combined probability of .39. However, the probability of the consistent changes of the percentages occurring in the expected direction is $1 \div 24$, and the product of the two probabilities is .016. This indicates that the increase of the number of RRRR cases with reading acceleration is probably not attributable to fluctuations of random sampling.

There were only a few LLLL cases, and therefore it is impossible to decide whether one should speak of reading inferiority of consistently right-sided individuals, or of ipsilaterals in general. Actually, the left-handed ipsilaterals tended to be poor readers in our material, but there were too few of them to establish this finding. The right-handed ipsilaterals were superior to the LRRR, the IRRR and to the boys who were not consistent right-handers, considered as a group. The subdivisions of the not consistently right-handed group were too small for any conclusions.

6) The frequencies of various kinds of orientation errors (i.e. O, R, T and I errors) differed only slightly in the groups of boys of different laterality type; the group differences were not statistically significant. The directions of the differences did not constitute any plausible pattern. There appeared to be no relationship between laterality type and proneness to any type of error in the group results. Numerous combinations of

laterality patterns were tried out. The results shown in Table III are typical. The table presents the mean frequencies of O, R, T and I errors made by the ipsilateral group, by all non-ipsilaterals, by impartially-sighting boys and by crossed laterals, i.e. the LRRR and RLLL groups combined.

TABLE III.—MEAN FREQUENCIES OF O, R, T, AND I ERRORS

Ipsilaterals				All Except Ipsilater				Impart. Sight.				Crossed Laterals			
O	R	T	I	O	R	T	I	O	R	T	I	O	R	T	S
.7	0	.1	0	1.3	.25	.1	.2	.5	.5	0	0	1.2	.2	.1	.2
3.7	1.3	.8	.6	3.9	1.2	1.1	.8	4.5	1.2	1.7	1.5	4.2	1	1	.9
4.7	2.1	2.6	.8	4.8	2.2	1.4	.8	3.8	2.2	1.6	.6	6.7	1.3	1.6	.8
1.5	0.0	1.3	.9	5.0	1.6	1.1	.8	4.2	1.5	1.3	0.9	5.2	1.4	.3	.4

The ipsilaterals make fewer O-errors than the boys with crossed hand and sighting preferences; however, all the critical ratios are considerably less than 2, and the impartially-sighting boys make still fewer O errors on three of the four reading skill levels than the ipsilateral group. No other differences are consistently formed on different levels of reading skill. Similar findings have been reported by Monroe.[15]

7) A few individual cases present interest, even though they cannot be considered to be more than suggestive.

The two boys who made by far the largest numbers of R-errors were both left-handed throwers and batters who wrote right-handed. This may be chance coincidence; but, since there were only five left-handers writing with the right hand, coincidence is not too probable.

One of these boys made only one O-error on the test, the other one between ten and eleven (depending on the interpretation of an ambiguous misreading of a word with multiple errors), the third largest number. Neither of the two boys making the two largest numbers of O-errors made many R-errors. An examination of the extreme cases does not indicate any close relationship between O and R errors. Some boys made strik-

ingly many R errors, some strikingly many O errors. Only one made an unusually large number of both.

All three boys making the largest numbers of O-errors exhibited crossed lateral preferences. The boy who made very many of both O- and R-errors was a retrained left-hander sighting with the right eye; the other two were right-handers sighting with the left eye. Whether the proneness towards large numbers of O-errors is associated with crossed laterality cannot be determined; coincidence is not too improbable, since cross laterality was very common in our material.

Two boys were found who were left-sided in sighting and throwing, right-sided in writing and batting. Both had originally been left-handed writers tending to produce mirror script.

DISCUSSION

1) Most of the findings obtained by Orton and cited by him in support of his theory were essentially duplicated in our material; and yet, no 'strephosymbolia' has been found.

The reversals involving the right and left directions were found in considerable numbers, and the O-reversals were found to be associated with reading retardation, not only with poor reading. But reversals involving the vertical dimension were also found. Orton, Monroe and others have also observed them, but it is not clear how their occurrence could be explained in terms of Orton's theory. An examination of the intercorrelations failed to indicate that the O- and the R-errors were particularly intimately connected.

Motor performances which ought to be indicative of weak or inconsistent lateral cerebral dominance seemed to be associated with retarded reading; but they were not found to be associated in any identifiable way with reversal errors in the horizontal direction as they should have been according to Orton's theory.

2) Neither association between weak unilateral cerebral dominance and reading retardation, nor Monroe's result of superior mirror-reading and mirror-writing of retarded readers, nor the occurrence of reversal errors involving the horizontal direction requires an explanation in terms of normally inactive mirror images of visual traces.

Evidence of increased incidence of weak or unstable cerebral dominance has been found in subjects exhibiting a variety of non-adaptive forms of behavior, e.g. stutterers,[1,6] speeders and reckless car drivers,[19] psychotics,[2] psychopaths, drug addicts and psychotic feebleminded,[18] intellectually subnormal and feebleminded children.[13] A conflict between memory traces of visual shapes and their mirror images cannot explain psychosis, drug addiction or reckless driving. Apparently, a disproportionate number of individuals with a central nervous system without clearly established unilateral dominance exhibit emotional and other difficulties; the nature of this relationship is by no means clear. There may be a primary physiological relation between laterality and emotional control, or the statistical relationship may be due to the frequently harsh measures taken by parents or teachers to ensure right-handedness in children.[4] Facts could be cited in support of both viewpoints, the second one having the slight advantage that it might explain the puzzling variations in the results of different investigators.

Emotional difficulties may or may not account for the apparent relationship between weak lateral cerebral dominance and reading difficulties found in our data. Such a relationship was also found in some other studies,[14,20,23] though not in all dealing with the problem. But, since indications of weak lateral cerebral dominance were found to be rather common not only among retarded readers, but also among psychopaths, reckless drivers, and drug addicts, there cannot be any necessity to explain any of these facts in terms of mirror images of memory traces.

The simplest explanation of the relative superiority of retarded readers in mirror-reading and mirror-writing is their superior intelligence, which should make them better able to master these new tasks. Retarded readers are necessarily superior in intelligence to children who match them in reading skill and whose intelligence is commensurate with their reading; in the retarded readers, the level of intelligence is by definition above the level of their reading skill.

Nor does the occurrence of reversal errors require the existence of mirror images of memory traces for its explanation. There are numerous findings (Orton quotes several) indicating that orientation is a rather subordinate aspect of visual patterns, and that it is relatively more subordinate in early childhood

than later on; thus confusion between shapes differing in orientation which occurs in retarded readers appears to be a plausible effect of the persistence of an infantile mode of behavior, possibly a phenomenon of the fixation type occurring in children with emotional difficulties. The following facts may be cited in support of this hypothesis:

It takes only a few days for a subject to feel comfortable in an inverted visual world.[3,21] If the orientation is changed to a slighter extent, e.g. by tilting the visual field by pressure on the eye-ball, the tilted appearance of objects disappears within a few minutes.[10] Koffka[9] found a number of published observations of children looking at picture books in upside-down position and readily recognizing the objects depicted. Fildes[3] reported that children with reading difficulties tend to have greater difficulties in remembering the orientation of visual shapes than normal children. Visual shapes differing only in orientation seem to be relatively equivalent in behavior, more so in preschool children and retarded readers.

No physiological explanation of the relative equivalence of shapes irrespective of orientation can be given. There are many related phenomena of 'equivalence of stimuli'[7,8,11] none of which can be adequately explained in physiological terms, but the true explanation of the sum total of them definitely cannot be of the type postulated by Orton for the equivalence of mirror images for memory in case of some poor readers. Nobody has ever suggested that every visual pattern leaves as many different separate memory traces as there are 'equivalent stimuli' for it, differing in color, size, orientation in different directions, etc.

3) The preceding discussion does not negate a possible relationship between the laterality patterns of occasional children and their reversal errors in reading. On the contrary, if shapes differing only in orientation are relatively equivalent for us, it is probable that kinesthetic cues should contribute to our discrimination between them; this idea has an old history in psychology, and Orton recognizes the probable rôle of kinesthetic factors in visual experience. It is to be expected that absence of clear-cut lateral motor preferences should occasionally lead to visual confusions of kinesthetic origin; however, our statistical findings, as those of others, indicate that cases of this type are not sufficiently

frequent to show up in moderately sized groups. However, the extreme cases discussed earlier are rather suggestive of this possibility. The three boys making many O errors, the two making R errors, the two with a history of spontaneous mirror-writing, exhibited three characteristic laterality patterns, a different one in each of these three cases. Van Riper[22] also reported a finding with ambidextrous subjects, suggestive of a relationship between visual and kinesthetic recall.

4) The objection might be raised that the intellectually subnormal group of subjects may not be comparable to normal groups. However, this objection is hardly valid, for two reasons. First, several of Orton's and Monroe's own results were duplicated, none of their findings contradicted. Second, the findings agree with those of several other investigators using normal subjects; e.g., Monroe in a later publication[15] also failed to find any significant relationship between laterality type and reversal frequency. L. S. Wolfe[15] did not find any unusual distribution of laterality types in her group of normal school children selected on the basis of proneness to reading reversals. The present writer has failed to find any contrary experimental evidence in the literature surveyed by him.

SUMMARY

The reading performance, handedness and sighting preferences of ninety-five intellectually subnormal boys were examined. The results were opposed to Orton's theory of a double set of memory traces of visual shapes and lack of unilateral cerebral dominance as a major factor in reading disability—there were too many types of reversal errors; the reversal errors involving confusion between right and left were not sufficiently intercorrelated; there was no demonstrable relationship in the group data between the laterality pattern and the error patterns of the subjects. Most of Orton's results were duplicated, which argues against the assumption that the reading performance of our subnormal subjects differed in essential features from that of the normal population. Numerous reversal errors were found; some types were correlated specifically with reading retardation and among the retarded readers the percentage of consistently right-sided subjects was rather low. Interpretations of the findings are offered.

REFERENCES

1) Bryngelson, B., "A study of laterality of stutterers and normal speakers," *J. Soc. Psych.*, 11, 1940, 151–155.

2) Cornell, C., "Eye, hand and foot preferences of psychotic patients compared with college students," *J. Juv. Research*, 22, 1938, 115–118.

3) Ewert, H., *A study of the effect of inverted retinal stimulation upon spatially coördinated behavior*, Genet. Psych. Monogr., 7, 1930, 177–363.

4) Fildes, L. C., "A psychological inquiry into the nature of the condition known as congenital word blindness," *Brain*, 44, 1921, 286–307.

5) Haefner, R., *The educational significance of lefthandedness*, Teachers College Contrib. to Ed., 1929, 360.

6) Jasper, H. H., *A laboratory study of diagnostic indices of bilateral neuromuscular organization of stutterers and normal speakers*, Psych. Monog. 43, 1932, 72–144.

7) Klüver, H., "The equivalence of stimuli in the behavior of monkeys," *J. Genet. Psych.*, 34, 1931.

8) Klüver, H., "The study of personality and the method of equivalent and non-equivalent stimuli," *Character and Personality*, 5, 1936, 91–112.

9) Koffka, K., *Growth of the Mind*, New York: Harcourt Brace & Co., 1931, 313–14.

10) Koffka, K., *Principles of Gestalt Psychology*, New York: Harcourt, Brace & Co., 1935, 215.

11) Kries, J. von, *Uber die materiellen Grundlagen der Bewusstseinserscheinungen*, Tübingen: 1901.

12) Miles, W. R., "Ocular dominance demonstrated by unconscious sighting," *J. Exper. Psych.*, 12, 1929, 113–126.

13) Mintz, A., "Lateral preferences in a group of intellectually subnormal boys," *J. Genet. Psych.* (accepted for publication).

14) Monroe, M., *Methods for diagnosis and treatment of reading disability*, Genet. psych. monog., 4, 1928, 343–456.

15) Monroe, M., *Children who cannot read*, University of Chicago Press, 1932.

16) Orton, S. T., "Word-blindness in school children," *Arch. Neurol. and Psychiat.*, 14, 1925, 581–615.

17) Orton, S. T., "Visual functions in strephosymbolia," *Arch. Ophthalm.*, 30, 1943, 707–717.

18) Quinan, C., "The principal sinistral types," *Arch. Neur. Psychiat*, 24, 1930, 35–47.

19) Quinan, C., "The handedness and eyedness of speeders and of reckless drivers," *Arch. neurol. psychiat.*, 25, 1931, 829–837.

20) Schonell, F. J., "The relation of reading disability to handedness and certain ocular factors, II," *Brit. J. Educ. Psych.*, 11, 1941, 20–27.

21) Stratton, C. M., "Vision without inversion of the retinal image," *Psych. Review*, 1897, 341–360.

22) Van Riper, "A new test of laterality," *J. Exper. Psych.*, 17, 1934, 305–313.

23) Wile, I. S., "Eye dominance," *Arch. Ophthalomol.*, 28, 1942, 780–790.

24) Wolfe, L. S., "Differential factors in specific reading disability, I, Laterality of function," *J. Genet. Psych.*, 58, 1941, 45–56.

NOTE ON THE ESTIMATION OF COLLEGE APTITUDE TEST SCORES FROM IQ'S DERIVED FROM GROUP INTELLIGENCE TESTS

R. B. EMBREE, JR.

College of William and Mary

During recent months, two articles have appeared in this JOURNAL which attempt to provide for the prediction of college aptitude test scores from intelligence quotients determined several years prior to graduation from high school. Edmund G. Weber[1] in the October 1945 issue presented an equation based upon the Otis Quick-Scoring Tests of Mental Ability and the American Council on Education Psychological Examination. In the March 1946 issue, David E. Votaw[2] offered a similar conversion making use of regression equations between Otis Group Intelligence Test IQ's and scores on the American Council Psychological Examination. It was suggested by both authors that such information may be of value to counselors in high school who wish to anticipate the standing of students on subsequent college aptitude examinations.

A careful reading of the two articles will indicate that the conversion processes appear to produce very divergent results. Weber's group had a mean IQ of 113.34 and a mean ACE raw score of 76.79. Votaw's group had a mean IQ of 108.9 and a mean ACE raw score of 92.4. The groups were quite similar with respect to the variability of IQ, but the standard deviation of ACE was greater for Weber's population. An IQ of 110 is equivalent to an ACE score of approximately 68 on Weber's graph and a score of approximately 95 on Votaw's regression chart. Obviously, the counselor will be perplexed if he happens to be familiar with both articles, since this difference appears to amount to more than a standard deviation unit in ACE score. The relative merits of linear equation and the regression tech-

[1] Weber, E. G. "Equating High-School Intelligence Quotients with College Aptitude Test Scores." *The Journal of Educational Psychology*, 36, 1945, 443–446.

[2] Votaw, D. F. "Regression Lines for Estimating Intelligence Quotients from American Council Examination Scores." *The Journal of Educational Psychology*, 37, 1946, 179–181.

nique can be omitted from this discussion, since both devices should produce relatively similar results at or near the mean when applied to similar data.

The crux of discrepancy lies in the probability that similar data were not used in the two studies, although it is impossible to determine this fact from the reports. Any such equations for the estimation of college aptitude test scores from childhood intelligence quotients, if they are to be generally useful, must take into account the following points:

1) The intelligence test used to determine the IQ's, the conditions under which it was given, and the age of subjects at the time of testing.

2) The type and form of college aptitude test employed and the conditions under which it was administered.

3) The correlation of IQ's with the subsequent college aptitude test.

4) The characteristics of the group test IQ, with special reference to its variability.

Both writers furnish information with regard to the intelligence tests used and the approximate ages of subjects at the time of examination. However, as is implied in point (4) above, generalizations regarding the significance, in terms of ACE scores, of IQ's derived from other group tests would have to be made with extreme caution.

Neither writer specified what form of the American Council Psychological Examination was employed in his study. It is likely that herein lies the major cause of discrepancy. Weber's graduates must have taken the 1937 Experimental Machine-Scoring edition of the examination which was used by the University of Minnesota during the time when his cases were tested. It is difficult to secure a reliable national norm for this form of the ACE. However, the mean score on this test of University of Minnesota freshmen in 1938 was 78, and the standard deviation was 23. Votaw's cases, entering college in 1943–1945, were probably examined on contemporary forms of the ACE. From 1941 through 1944, the national college freshman means determined from various forms of the ACE were near 105, with standard deviations near 26.

The American Council Psychological Examination is perhaps the most widely used college aptitude test, and a different form is issued each year. Since 1941, when the examination was established in a single form for both machine- and hand-scoring, the authors have endeavored to provide annual editions of equal difficulty. Nevertheless, studies combining the results of these tests over a period of years must be interpreted with caution, since the annual tables of equivalent scores have been based upon the assumption that the freshman populations taking the tests from year to year were equal in terms of abilities measured by the examination.

BOOK REVIEWS

W. A. Brownell (Chairman), H. R. Douglass, W. G. Findley, V. M. Sims, and H. F. Spitzer (Society's Committee). *The Measurement of Understanding.* Forty-fifth Yearbook, Part I, National Society for the Study of Education, 1946, pp. 338. (paper)

The Measurement of Understanding is significant. The editor's avowal that the work is designed "as an aid to classroom teachers in the improvement of evaluative procedures" is capable of fulfillment. It is much more than a collection of suggestions and illustrations for evaluation purposes. The teacher's first task is to bring the student to an understanding; measurement of understanding is secondary, except as it is envisaged as an integral part of the teaching of understanding. This volume should be an aid to the teaching of understanding as well as to its measurement. It is not for the expert in measurement, but for the teacher.

Chapter I presents the purpose and scope of the Yearbook. Section I (Chapters II–IV) treats "The Importance of Teaching for Understanding," "The Nature of Understanding," and "Obtaining Evidence of Understanding." Section II contains discussions and many helpful illustrative methods of measuring—at least of obtaining evidence of—understanding in the following subject-matter areas (by way of a chapter devoted to each): social studies; science; mathematics, elementary; mathematics, secondary; language; fine arts, health education. physical education; home economics; agriculture; technical education; industrial arts. Section III presents a critical 'Forward Look.'

The chapters of Section II vary in quality, as is to be expected. In the writer's opinion, they vary from fairly good to excellent. Certainly it will occur to the reader that it is much easier to do an excellent job in some subject-matter areas than in others. Each of the chapters starts with a listing of the recognized teaching objectives. In the main, these are taken from pronouncements of committees of national organizations or other societies whose aim has been to foster good teaching in the various areas. Subsequently, testing procedures and other means of evaluating meaningful learning with respect to each of the objectives are presented. Subject to limitations in evaluative procedures the methods of measuring understanding should be as valid as

the objectives. The committee is to be congratulated upon the securing of this synthesis of measurement and teaching objectives.

The authors have made liberal use of techniques other than objective tests. In some chapters, for example the fine arts, objective tests are scarcely used at all. In others, for example science, considerable use is made of the better type of objective tests. In all, transfer—by which is here meant the ability to apply information, principles, procedures, etc., learned in one situation in attacking other relevant situations—has been one of the guiding principles.

This volume should have wide use in professional courses in education. It should also prove helpful to classroom teachers generally.

J. B. STROUD
State University of Iowa

ELIZABETH STEVENSON. *Home and Family Life Education in Elementary Schools.* New York: John Wiley and Sons, Inc. 1946. pp. 309.

This book is addressed to the entire educational community in an effort to promote more and better education for 'personal, home, and family living.' Part I, "A New Concept of Home and Family Life Education," develops a conceptual framework for understanding Part II, and includes discussions of the present status of this area in elementary education, the developmental needs of children and their relation to the psychology of learning, the school's responsibility in a democratic society and the relationship of this program to the goals of general education.

In Part II, "Suggested Experiences in Personal, Home and Family Life," the author brings together a rich variety of experiences which have been found useful in accomplishing some of the objectives stated in Part I. Consideration is given to such units as "learning to live with others," "food for pleasure, health and sociability," "pupils' responsibilities to younger children and convalescents in the home," "making home a better place in which to live," "learning how to dress," and "becoming an intelligent consumer."

For many readers Part II will be the most valuable part of the book, for the specific experiences are not merely described, but are interwoven within a philosophy of education which puts considerable emphasis on the importance of beginning with the

immediate problems of the learner. Attention is also given to the importance of (a) developing from specific experiences larger generalizations which might influence behavior in a wider variety of situations and (b) continuously evaluating procedures.

Aside from the contribution Miss Stevenson makes by suggesting to teachers experiences for implementing their own programs, the value of her book must be judged by the defense of the theses developed in Part I. Essentially the volume pleads for a neglected area of educational experience. The author's concern is reflected in the statement that "education for personal, home and family living should be part of each person's education" (p. 28). The reviewer is favorably disposed toward a point of view which stresses the need for better integration of family and school experiences and which recognizes the impact of intra-family behavior on the school activities of children. However, there is always the danger that recognition of social needs can easily lead to an imposition of adult-felt necessities on children.

The author may have reconciled to her own satisfaction a desire to promote better 'personal, home and family life education' with her conviction that the immediate needs and problems of children must be the basis of educative experience, but the reviewer got the impression that she used good educational psychology (starting where the learner is, recognizing the importance of first-hand perceptual experience for learning) to accomplish goals considered by adults to be desirable for children.

The author is sensitive to ways in which immediate experiences provide the locus for building more inclusive concepts. "Often the child is unable to see the relationship of the national, the community or even the neighborhood life to his personal life. His understanding and interest might be increased by approaching the objective world through a personal need." Although written in a context considering family diets, this has interesting implications for those seeking a world community overnight.

The many important insights contained in this book cannot obscure the weakness of the theoretical presentation in Part I. Topics such as the meaning of democratic experiencing and the continuity of experiences are treated too sketchily to be convincing to the skeptic, while the reader who already thinks the author's way will find only affirmation.

In bringing together a variety of experiences found helpful by others, in assembling an excellent bibliography for use of teachers and students, the author has contributed to an area of education too often neglected. For the reader in search of significant thinking about basic problems in elementary education the book has less to offer.

PAUL E. EISERER
The University of Chicago

FREDERICK L. WESTOVER. *Controlled Eye Movements versus Practice Exercises in Reading.* New York: Bureau of Publications, Teachers College, Columbia University, 1946, pp. 99.

Several mechanical devices are now available for controlling eye movements in reading. Their use, which has become widespread, is based upon the assumption that the pacing of eye movements will improve reading proficiency by correcting faulty oculomotor habits. The technique, unfortunately, emphasizes peripheral factors as fundamental in causing reading disability. But the use of pacing techniques, with or without the aid of mechanical devices, have quite uniformly produced improvement in the rate of reading. Although a few studies have been concerned with the effectiveness of pacing techniques in comparison with other methods of improving reading, more evidence is needed. This report deals with a comparison of three methods of improving the reading speed and comprehension of college freshmen: (1) college work with no special exercises although members of this (the control group) as well as the other groups were motivated by informing them that their reading ability was poor, that they needed to do better to handle college work and that improvement was possible; (2) college work with practice in special reading exercises; (3) college work with practice in reading the same special exercises under conditions of controlled eye movements (pacing).

There were one hundred forty 'poor' college subjects (poor readers) divided into three matched groups. The two experimental groups received five weeks of training, two hours per week. The control group received no formal training in reading. Initial tests were administered to all subjects. At the end of the training, the reading tests were repeated and were given again six months after the training was completed. A special appara-

tus was devised to present the reading exercises in a controlled manner (pacing).

All three groups made significant gains in reading achievement. Gains in rate of reading were significantly greater for the two experimental groups than for the control group, but there were no significant differences in the effectiveness of the two instructional methods (reading exercises vs. pacing). At the close of the term there were no significant differences in the grade points earned by the three groups, but the two groups receiving special training in reading achieved better scholarship and had fewer withdrawals from college than the control group. Further gains were found after six months, but these gains in reading speed and comprehension were not significantly different for the three groups. The finding that mechanical pacing of eye movements yields results no different from that achieved by the use of reading exercises is highly important. It gives additional experimental support to the view frequently expressed by the reviewer that just as satisfactory gains in reading may be achieved without the use of pacing techniques. When pacing techniques are employed, the teacher or the clinical worker are too prone to overemphasize peripheral factors in reading and to neglect the more fundamental central factors of perception, comprehension and assimilation.

This investigation was adequately designed and the data skillfully analyzed. The author is well aware of certain limitations of the investigation, i.e., the training period was rather short; the 'control group' was not a real control group since its members were motivated to improve their reading. This excellent study will be well received by those interested in the experimental study of reading, but will be ignored by those teachers and clinicians who love their mechanical gadgets.

MILES A. TINKER

University of Minnesota

STEWART G. COLE, I. JAMES QUILLEN, AND MILDRED J. WIESE. *Charting Intercultural Education*, 1945–55. Stanford University: Stanford University Press, 1946. pp. vi + 58.

This booklet is a report of the Summer 1945 Workshop on Intercultural Education conducted at Stanford University. Its purpose is "to provide a brief guide for teachers, administrators, and

lay leaders . . . interested in the improvement of interracial and intercultural relations." It makes a plea for education designed to foster participation as good citizens in mixed communities and in national and international affairs. To the realization of this end are proposed the dissemination of scientific information about race, inequality of opportunity, requirements of good living by all citizens; working for the enactment of legislation forbidding discriminatory practices; cultivation of neighborliness and all-American citizenship.

Approximately half of the space is devoted to specific and practical suggestions for implementing the program, in (1) the elementary school, (2) the secondary school, and (3) the community. Then are suggested materials for study.

The booklet is written in persuasive style and the practical suggestions are well considered. It is timely and should be effective.

J. B. STROUD

State University of Iowa

PUBLICATIONS RECEIVED

JADWIGA ABRAMSON. *L'Enfant et L'Adolescent Instables, Etudes Cliniques et Psychologiques.* Paris: Presses Universitaires de France, 1945, pp. 390. (paper)

RICHARD D. ALTICK. *Preface to Critical Thinking.* New York: Henry Holt & Co., 1946, pp. 321. $1.60.

PAUL BODIN. *L'Adaptation de L'Enfant au Milieu Scolaire.* Paris: Presses Universitaires de France, 1945, pp. 156. (paper)

BENJAMIN, C. BOWKER. *Out of Uniform.* New York: W. W. Norton and Co., 1946, pp. 259. $2.75.

FRANK A. BUTLER. *Improvement of Teaching in Secondary Schools.* Revised Edition. Chicago: The University of Chicago Press, 1946, pp. 399. $3.50.

RAYMOND B. CATTELL. *Description and Measurement of Personality.* Yonkers, N. Y.: World Book Co., 1946, pp. 602.

FREDERICK B. DAVIS. *Item-Analysis Data: Their Computation, Interpretation, and Use in Test Construction.* Cambridge, Mass.: Graduate School of Education, Harvard University, 1946, pp. 42. (paper) $0.75.

LEONHARD FRANK. *Dream Mates.* Translated by Maxim Newmark. New York: Philosophical Library, 1946, pp. 245. $3.00.

BERTHA B. FRIEDMAN. *Foundations of the Measurement of Values. The Methodology of Location and Quantification.* New York: Bureau of Publications, Teachers College, Columbia University, 1946, pp. 227. $2.75.

BENJAMIN C. GRUENBERG, EMILY EVELETH SNYDER, AND JESSE V. MILLER. *Workbook for Students of Biology.* Boston: Ginn & Co., 1946, pp. 267. (paper) $1.32.

PHILIP LAWRENCE HARRIMAN, Editor. *Encyclopedia of Psychology.* New York: Philosophical Library, 1946, pp. 897. $10.00.

RAYMOND H. HUGHES AND WILLIAM H. LANCELOT. *Education—America's Magic.* Ames, Iowa: The Iowa State College Press, 1946, pp. 189. $2.50.

HOWARD L. KINGSLEY. *The Nature and Conditions of Learning.* New York: Prentice-Hall, Inc., 1946, pp. 579. $4.50.

DARREL J. MASE. *Etiology of Articulatory Speech Defects.* New York: Bureau of Publications, Teachers College, Columbia University, 1946, pp. 85. $2.10.

THEODOR REIK. *The Psychological Problems of Religion. I: Ritual—Psychoanalytic Studies.* New York: Farrar, Straus and Co., 1946, pp. 367. $5.00.

T. W. RICHARDS. *Modern Clinical Psychology.* New York: McGraw-Hill Book Co., 1946, pp. 331. $3.50.

HORACIO J. A. RIMOLDI. *Ritmo y Fatiga: Problemas que plantea la hipotesis de la e energia mental.* Buenos Aires: Libreria y Editorial "El Ateneo," 1946, pp. 190. (paper)

HELEN M. ROBINSON. *Why Pupils Fail in Reading.* Chicago: The University of Chicago Press, 1946, pp. 257. $3.00.

S. R. SLAVSON. *Recreation and the Total Personality.* New York: Association Press, 1946, pp. 205. $3.00.

JOSEPH TIFFIN, FREDERIC B. KNIGHT, AND ESTON JACKSON ASHER. *The Psychology of Normal People.* Revised Edition. Boston: D. C. Heath & Co., 1946, pp. 581.

ARTHUR E. TRAXLER AND AGATHA TOWNSEND. *Another Five Years of Research in Reading: Summary and Bibliography.* New York: Educational Records Bureau, 1946, pp. 192. (paper)

P. F. VALENTINE, Editor. *Twentieth Century Education: Recent Developments in American Education.* New York: Philosophical Library, 1946, pp. 655. $7.50.

HENRI WALLON. *Les Origines de la Pensee Chez L'Enfant.* Volume I: "Les Moyens Intellectuels," pp. 307. Volume II: "Les Taches Intellectuelles," pp. 449. Paris: Presses Universitaires de France, 1945. (paper)

ROGER J. WILLIAMS. *The Human Frontier: A New Pathway for Science toward a Better Understanding of Ourselves.* New York: Harcourt, Brace & Co., 1946, pp. 314. $3.00.